Aalto University
School of Science
Department of Mathematics and Systems Analysis

Antti Levo

Predicting Pilot Fatigue in Commercial Air Transportation

Master's Thesis
Espoo, January 15, 2016

Supervisor: Prof. Ahti Salo
Thesis advisor: M. Sc. (Tech.) Janne Kulmala

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Antti Levo

Title of thesis Predicting Pilot Fatigue in Commercial Air Transportation

Degree programme Engineering Physics and Mathematics

Thesis supervisor Prof. Ahti Salo **Code of professorship** Mat-2

Department Department of Mathematics and Systems Analysis

Thesis advisor(s) M. Sc. (Tech) Janne Kulmala

Date 15.01.2016 **Number of pages** 5+74 **Language** English

Abstract

Fatigue is a major human factor related cause of aviation accidents and currently subject to increased discussion by aviation authorities and professional pilots. Reducing fatigue and minimizing the risk of fatigue induced errors requires predicting the alertness state of crew members and identifying possible fatigue sources. Bio-mathematical models are a way to estimate fatigue levels of crew members based on work schedules. These models are currently utilized in commercial aviation to various degrees.

In this thesis a prediction model for estimating fatigue is developed. It is based on theory of sleep and fatigue and previous research on bio-mathematical models. Fatigue data was gathered from pilots working on short-haul operations and it is used to estimate parameters and validate the model. Work schedules are used as input and an alertness score based on Karolinska Sleepiness Scale is estimated. The aim is to estimate fatigue in the work schedule planning phase, weeks in advance of actual date of operations. The sources for fatigue in the model are time of day, time worked and presence of consecutive early morning shifts. Time not in work is defined as recovery period, which decreases fatigue. Cumulative effects were not identified to have significant effect on fatigue with available data.

The results indicate that it is possible to develop a model that estimates fatigue adequately, but the personal differences how people feel and experience fatigue make it difficult to create an applicable average model that fits well for everyone. Estimating parameters for every individual increases the accuracy and makes the model more feasible, but that is not practical for extensive use in flight operations due to data and time requirements. The developed model is, however, usable as a risk management tool in order to identify fatigue hazards, but it cannot be used as a sole basis for decision making due to limited accuracy. The greatest problem is the lack of sleep data as the amount of sleep has major impact on fatigue levels. Models used to estimate fatigue in advance need to either estimate the amount of sleep based on some probabilistic method or omit it from the model completely. In this thesis sleep was not included in the model.

Keywords Fatigue, Bio-mathematical model, Fatigue risk management system

Aalto-yliopisto
Perustieteiden korkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Diplomityön tiivistelmä

Tekijä Antti Levo	
Työn nimi Lentäjien väsymystilan ennustaminen kaupallisessa lentotyössä	
Koulutusohjelma Teknillinen fysiikka ja matematiikka	
Valvoja Prof. Ahti Salo	**Professuurikoodi** Mat-2
Laitos Matematiikan ja systeemianalyysin laitos	
Työn ohjaaja(t) DI Janne Kulmala	
Päivämäärä 15.01.2016 **Sivumäärä** 5+74	**Kieli** Englanti

Tiivistelmä

Väsymys on merkittävä inhimillinen tekijä ilmailuonnettomuuksissa ja se on tällä hetkellä lisääntyneen tarkastelun alla viranomaisten ja lentäjien keskuudessa. Väsymyksen ja sen aiheuttamien riskien minimointi vaatii lentomiehistön väsymystilan ennustamista ja mahdollisten lähteiden löytämistä. Biomatemaattiset mallit ovat keino arvioida miehistön väsymystilaa perustuen työvuorolistoihin ja niitä käytetään tällä hetkellä kaupallisessa ilmailussa vaihtelevissa määrin.

Tässä diplomityössä muodostetaan malli ennustamaan väsymystilaa. Malli perustuu uni- ja väsymysteorioihin ja se hyödyntää edeltävää tutkimusta biomatemaattisista malleista. Lyhyen kantaman lentojen lentäjiltä on kerätty dataa heidän väsymystiloistaan ja tätä dataa käytetään apuna mallin parametrien arvioinnissa. Malli käyttää syöttönä työvuorolistoja ja arvio väsymystilaa Karolinskan väsymysasteikolla. Tarkoitus on ennustaa väsymystä työvuorosuunnitteluvaiheessa viikkoja ennen todellista operointipäivää. Mallissa väsymystä aiheuttavina tekijöinä käytetään vuorokauden aikaa, työskentelyaikaa sekä aikaisia aamuherätyksiä. Työn ulkopuolisen ajan oletetaan olevan palautumisvaihetta, jolloin vireys kasvaa. Datan perusteella kumulatiivisten vaikutusten ei huomattu olleen merkittäviä.

Tulosten perusteella on mahdollista muodostaa malli, joka ennustaa tyydyttävästi väsymystä. Henkilökohtaisten erojen takia on kuitenkin vaikea muodostaa koettua väsymystilaa kuvaava malli, joka sopii jokaisen eri ihmisen väsymystilan ennustamiseen. Parametrien määrittäminen erikseen kullekin henkilölle kasvattaa mallin tarkkuutta, mutta on samalla työlästä ja aikaa vievää, eikä siten sovellu laajaan käyttöön lento-operoinnissa. Kehitetty malli toimii kuitenkin työkaluna riskien hallintaan ja väsymyksestä johtuvien vaaratilanteiden ennustamiseen. Sitä ei voi käyttää ainoana päätöksentekovälineenä rajoittuneen tarkkuuden takia. Suurin ongelma mallin kohdalla on uneen liittyvän datan puute, koska unella on merkittävä vaikutus väsymystilaan. Käytettäessä mallia, joka ennustaa väsymystilaa pitkälle tulevaisuuteen, unen määrä täytyy estimoida jollain todennäköisyysfunktiolla tai jättää pois mallista. Tässä työssä unen määrä ja ajoittuminen on jätetty mallin ulkopuolelle.

Avainsanat Väsymys, Biomatemaattinen malli

Contents

Abbreviations

AIP	Aeronautical Information Publication
CAA	Civil Aviation Authority
EASA	European Aviation Safety Agency
ECA	European Cockpit Association
FDP	Flight Duty Period
FRMS	Fatigue Risk Management System
FTL	Flight Time Limitation
ICAO	International Civil Aviation Organization
KSS	Karolinska Sleepiness Scale
PVT	Psychomotor Vigilance Task
REM	Rapid Eye Movement
SMS	Safety Management System
SP	Samn-Perelli

Foreword

Finishing this thesis has been a long, but rewarding journey. It took one changing of the subject, which turned out to be the right decision in the end. I’m happy to have been able to combine both my studies in engineering and aviation when writing this work, which made the creation process much more interesting.

I would like to thank my supervisor and superior Janne Kulmala for giving me this chance and offering helpful guidance towards completion. I wish also to thank Professor Ahti Salo for overseeing my thesis and giving me valuable feedback and showing interest to my topic.

I want to thank my family and friends for being there for me; my parents for supporting me through life and my brothers for growing and learning together.

Finally, I would like to thank my girlfriend and most important person in my life, Julia, for supporting me with writing this thesis and being there for me in the ups and downs of life.

Espoo, January 15, 2016
Antti Levo

1 Introduction

1.1. Background

Safety has been the number one priority in aviation throughout its history and the efforts are continuing to reduce and minimize the risk of incidents. Fatigue is a major human factors risk especially in commercial aviation, because of the shift work nature and the direct effects to pilots' abilities to perform their job safely (Caldwell et al. 2009). Competition between airline companies is intense, and consequently crew members are often pushed to their limits. Changes in a larger scale can be achieved through good legislation and industry wide standards, because there will otherwise always be companies that cut corners in order to save costs.

The technological advances over the 100 years of aviation history have been dramatic, while the human physiology has not changed at all. There are limitations to the ability of human body that cannot be overcome. The present need is for 24-hour operations, with around the world flights that cross multiple time zones, while having irregular and sometimes unpredictable work schedules. These factors lead to performance impairing fatigue, which poses a great risk to safety. In a barometer assembled by the European Cockpit Association (ECA) from more than 6000 European pilots, alarming results were found. Over 50 % of surveyed pilots experienced fatigue, which impaired their ability to perform while on duty, and 4 out of 5 pilots felt that they have to cope with fatigue while in cockpit (ECA, 2012)

Fatigue is currently managed with limitations to crew duty and rest times. These rules are used to protect crew members from hazardously low levels of awareness. Airline companies try to utilize crew to its maximum usability, which causes possible conflicts with cost and safety. Goode (2003) presents findings of increased probability of an accident the greater the hours of duty time for pilots. There have also been several accidents in the last 20 years with proof of fatigue relating to the incident. This has alerted industry leaders to the major safety risks caused by fatigue. Aviation as an industry becomes more mature from the mistakes that are made in order to avoid them in the future.

Although the understanding of fatigue, sleep and physiology has advanced over the past decades, current regulations have not been able to incorporate the new knowledge (Caldwell et al. 2009). Scientists and flight crew in cooperation are creating bio-mathematical models in order to better understand and predict fatigue and alertness. Based on the first models developed in the 80's (Borbely, 1982), these bio-mathematical models try to predict the probable fatigue of a crew member and are used in collaboration with risk analysis tools to identify possible problems in operations. Because the fatigue risks cannot be completely eliminated, they must be managed well to ensure adequate safety levels. Utilizing scientific knowledge to manage fatigue promotes crew alertness and performance in operations and increases safety, which is the ultimate goal of the aviation culture.

1.2. Research Objectives

The aim of this thesis is to study fatigue risk modeling in commercial aviation regarding pilot fatigue in everyday operations. The case company is a regional, short haul operator, with single type aircraft. The flight crew consists of two pilots, a captain and a first officer, with no third pilot available to relieve them during flight.

Study objectives are to:

1. *Create a model to predict fatigue based on flight crew work rosters.*
2. *Validate the model through empirical data obtained from pilots.*

The model will be used to predict awareness level of pilots with inputs coming from work schedules and output being an alertness score. The model is validated with data gathered from pilots with self-assessment questionnaires. Extensive tests, such as brainwave measurements, are not practical in the scope of this thesis. The questionnaires are performed via iPad based forms to facilitate the collection of data without affecting everyday operations too much. A group of pilots is selected to work as a subject group in order to ensure adequate collection of information. Consistent data is required to validate the model in relevant situations and a test group that has been well briefed will allow the best results. The relevant inputs for the model are selected by researching literature and discussing the possible fatigue sources with experienced flight captains.

This thesis is limited to the creation of a prediction model. In the case of risk analysis this model is just a part of it. A thorough risk assessment requires further work, which is not part of this thesis project. The model may, however, be used as a part of a future risk management tool.

1.3. Structure

The structure in this thesis is formed into 6 chapters. First chapter introduces the topic, the goals and setup of this thesis. Chapter 2 gives background information of aviation industry and fatigue risk management. Chapter 3 presents the theory behind fatigue modeling. Chapter 4 presents the created fatigue model and the empirical data used to validate it. In chapter 5 the results are introduced and analyzed. Final chapter 6 presents the conclusion of this thesis.

2 Aviation Industry

2.1. Regulations

2.1.1. ICAO

Aviation is inherently highly international, which was recognized early from the start of the business. In 1944 the International Civil Aviation Organization (ICAO) was founded, operating as a specialized agency under the United Nations. The purpose of ICAO is to set the standards and recommended practices (SARPs) in international air traffic, which are presented in 19 Annexes to the Chicago Convention. Chicago Convention refers to the document signed by 52 signatory states in 1944, which established the ICAO and charged it with coordinating and regulating international air travel.

In its core, aviation is based on international agreements. These can be between two individual states, called bilateral agreements, or between several parties, such as with ICAO. Recognized all over the world, it is one of the fundamental multilateral agreements, with 191 member states joining it. The standardization of aviation industry has enabled the creation of global air transport network that can operate close to 100 000 daily flights. As an example of standardization, each member country should have an Aeronautical Information Publication (AIP) available, which contains information essential to air navigation. This AIP is updated every 28 days and provides airlines information about the airspace and aerodromes, allowing safe operations between and inside countries. (ICAO, 2014)

2.1.2. Civil aviation authorities

Each country in the world should have its own civil aviation authority (CAA). In Finland this position is governed by Trafi, Finnish Transport Safety Agency. This authority ensures that the aviation legislation follows the 19 Annexes of the Chicago Convention. In Europe the regulatory and executive task of civil aviation safety is governed by European Aviation Safety Agency (EASA). It collaborates with the national aviation authorities and is responsible for setting standards across EU. National authorities are required to implement EASA norms and regulations into their own regulations, which can be done by following SARPs or acquiring acceptance from EASA to national practice. Military and government aviation are however governed by national legislation. The CAAs oversee the aviation field in their own jurisdiction. They grant and supervise licenses and certificates, enforce regulations, safeguard passenger rights, maintain aviation registers, participate in national and international cooperation, deal with environmental issues related to aviation and provide advice and information. (Trafi [a], 2015)

2.2. Operations

Flight operations can be divided into four different categories depending on the duration or length of flights: short-, medium-, long- and ultra long-haul. There are no agreed definitions for these four categories, but they can be defined so that short-haul is flights with less than 3 hour travel, usually domestic flights, medium-haul between 3 to 7 hours, long-haul over 7 hours and ultra long-haul is flights with over 16 hour travel time. Increasing flight times increases the number of crew needed. Short-haul operations are carried with two pilots and paired in conjunction to create a daily work schedule. Ultra long-haul might instead require up to 4 pilots per flight to satisfy fatigue regulations and pilots may only operate a few given flights in a month in order not to exceed monthly work hour limitations. (ICAO, 2014)

The nature of operations defines the need of supporting functions. EASA regulations define that every aircraft operated under operator's certificate must be under continuous operational control, which means that an operations centre must be manned whenever there are operations on-going (EASA, 2012). Operators with only short-haul flights will not usually need 24-hour operational capability, as the production is not continuous over night. Long-haul operations, however, require constant presence in the operations centre, because there is usually always aircrafts in air throughout a day. Aircrafts on duty require flight monitoring and support regarding maintenance, crew rotation, flight planning and passenger flow. Without assistance operations cannot continue smoothly when problems are faced.

Work schedules for flight crew members with different types of operations are also very different. Intercontinental flights result in crew layovers in different time zones other than home base, whereas short-haul flights keep the crew in practically the same time zone every day. Nowadays rotation of aircraft is done so that crew does not normally have to stay idle away from home for too long. Before the current expansion of air travelling flight crew might have staid several days in hotels waiting for aircrafts return trip, as the flight networks were smaller and daily flights fewer, especially in long-haul traffic. In short-haul the utilization of crew is more flexible, as the distances are smaller and aircraft rotations faster.

2.3. Fatigue Management

2.3.1. Fatigue risk management system

Crew member fatigue can be defined as follows (ICAO, 2011):

A physiological state of reduced mental or physical performance capability resulting from sleep loss or extended wakefulness, circadian phase, or workload (mental and/or physical activity) that can impair a crew member's alertness and ability to safely operate an aircraft or perform safety related duties.

A Fatigue Risk Management System (FRMS) is defined as follows (ICAO, 2011):

A data-driven means of continuously monitoring and managing fatigue-related safety risks, based upon scientific principles and knowledge as well as operational experience that aims to ensure relevant personnel are performing at adequate levels of alertness.

The aim of the FRMS is to ensure that crew members are well rested and alert enough to safely operate aircraft without endangering passengers. It is a way to systematically manage the risks related to fatigue, balancing between productivity, costs, and safety. Principles and processes from Safety Management System (SMS) are applied to the FRMS to assure safety. In ICAO Annex 19 SMS is defined as a systematic approach to managing safety, including the necessary organizational structures, accountabilities, policies and procedures (ICAO, 2013). It is a formal risk management process that aims to identify, assess and mitigate risks. EASA regulations require operators to have both a safety risk management system (EASA, 2012), which is fulfilled by utilizing SMS principles, and a FRMS (EASA, 2014).

Through an effective safety reporting culture, both SMS and FRMS rely on the operating personnel to report hazards when observed. In order for both SMS and FRMS to work correctly, an operator must clearly distinguish between deliberate errors and unintentional human errors. This promotes a reporting culture where the flight crew will report events and issues without fear of punishment. In order to promote good safety culture in aviation, unintentional human errors should not be punished but need to be seen as possible situations for improvement. The scope and quality of a utilized FRMS enables airline operators to deviate from existing limitations and bring more flexibility into operations.

The ICAO requirements for FRMS processes are listed in Annex 6, part I, Appendix 8 (ICAO, 2010). They include the identification of hazards, risk assessment and risk mitigation, with the identification of hazards including predictive, proactive and reactive phases. Methods for predicting these hazards may include, but are not limited to, operational experience and data collected, evidence-based scheduling practices and bio-mathematical models. Notable here is that the bio-mathematical models are not a requirement but an approved method for predicting possible hazards of fatigue. The current limitations of the models are acknowledged and they may not be used alone to justify scheduling decisions. In general, they are only a minor part of the whole FRMS implementation process.

More important to the successful implementation of a FRMS is the collection of data and experience. The data required to apply a FRMS includes measuring fatigue levels of crew members and operational performance of the company. In addition to collecting data, this data must be analyzed to inform decisions made based on a FRMS. Through self-reporting, fatigue surveys, crew performance, scientific studies and analysis of planned versus actual time worked, proactive measures can be taken to identify and prepare for fatigue risks. All this requires collaboration from flight crew members and shared responsibility with managers and all involved personnel.

After preparing for risks, it is important to also analyze and react to the outcome of operations. Analyzing reports and incidents provides valuable information to better understand how fatigue related issues develop and what could be done differently. Determining whether a person was in a fatigued state is difficult and persons responsible for this must rely on information based on recall of the people involved.

2.3.2. Flight time limitations

Traditionally the approach to manage crew fatigue has been through flight time limitations (FTL), which are defined in ICAO Annex 6 that standardizes aircraft operations, maintenance and general aviation. EASA legislation (EASA, 2014) sets rules for flight and duty time limitations that follow the Annex 6 requirements. Air operators are required to follow these rules and to implement them in their crew work schedule planning. The purpose is to ensure that safety is not jeopardized by securing enough rest and limiting the amount of work of the flight crew. The FTL are simple rules that protect flight crew against fatigue.

New EASA legislation is currently being implemented. It will bring new changes to the FTL, which are aimed to increase crew alertness by reducing workload and restricting flight duty times. After the implementation, Europe will have one of the strictest FTL rules in the world (EASA, 2013). FTL rules, however, are only strict rules that do not consider fatigue on isolated operational level. They give requirements for minimum breaks and set maximum limits for daily, monthly, and yearly flight hours. For example, from EASA regulation 83/2014 (EASA, 2014)

PART.ORO.FTL.210

a) *The total duty periods to which a crew member may be assigned shall not exceed:*
   1) *60 duty hours in any 7 consecutive days;*
   2) *110 duty hours in any 14 consecutive days; and*
   3) *190 duty hours in any 28 consecutive days, spread as evenly as practicable throughout that period.*
b) *The total flight time of the sectors on which an individual crew member is assigned as an operating crew member shall not exceed:*
   1) *100 hours of flight time in any 28 consecutive days;*
   2) *900 hours of flight time in any calendar year; and*
   3) *1 000 hours of flight time in any 12 consecutive calendar months.*

A flight duty of 12:00 hours in some cases is allowed, but a duty of 12:01 hours is illegal. When estimating fatigue, the difference is insignificant. The rules are a bureaucratic way to protect crew against fatigue. They are same for all and do not consider major differences in operations, e.g. long haul versus short haul flights. They also consider linear relationship with working hours and fatigue (Steiner et al. 2012). The FRMS acknowledges all this and steers the focus to adequate alertness, which is more situational than the FTL rules consider. By considering fatigue more on the operational level and accounting for differences in flight routes, time of day and individual crew members, the operational flexibility can be increased while maintaining safety levels or even improving them.

EASA regulations (EASA, 2014) require authorities to decide whether operators under their jurisdiction belong to “early type” or “late type” when regarding disruptive schedules. Considering start of duties, early type of disruptive schedule means duties starting between 5:00 and 5:59 in the time zone to which a crew member is acclimatized. Late type means duties starting between 5:00 and 06:59. Finnish authority has decided to apply "early type" to all operators under its oversight (Trafi [b], 2015).

Time zone changes cause crew members to be adjusted to different times than local time where duty period should start. Acclimatized means a state in which a crew member's internal clock is synchronized to the local time zone. EASA regulation 83/2014 (EASA, 2014) considers crew members to be acclimatized to local time zone according to Table 1.

Table 1. State of acclimatization of crew members in time zone changes.

Time difference (h) between reference time and local time where the crew member starts the next duty	**Time elapsed since reporting at reference time**				
	< 48	48–71:59	72–95:59	96–119:59	≥ 120
< 4	B	D	D	D	D
≤ 6	B	X	D	D	D
≤ 9	B	X	X	D	D
≤ 12	B	X	X	X	D

- "B" means acclimatized to the local time of the departure time zone,
- "D" means acclimatized to the local time where the crew member starts his/her next duty, and
- "X" means that a crew member is in an unknown state of acclimatization

Maximum daily flight duty period (FDP) without extensions for acclimatized crew members is according to Table 2.

Table 2. Maximum daily FDP - Acclimatized crew members.

Start of FDP at reference time	1–2 Sectors	3 Sectors	4 Sectors	5 Sectors	6 Sectors	7 Sectors	8 Sectors	9 Sectors	10 Sectors
0600–1329	13:00	12:30	12:00	11:30	11:00	10:30	10:00	9:30	9:00
1330–1359	12:45	12:15	11:45	11:15	10:45	10:15	9:45	9:15	9:00
1400–1429	12:30	12:00	11:30	11:00	10:30	10:00	9:30	9:00	9:00
1430–1459	12:15	11:45	11:15	10:45	10:15	9:45	9:15	9:00	9:00
1500–1529	12:00	11:30	11:00	10:30	10:00	9:30	9:00	9:00	9:00
1530–1559	11:45	11:15	10:45	10:15	9:45	9:15	9:00	9:00	9:00
1530–1559	11:45	11:15	10:45	10:15	9:45	9:15	9:00	9:00	9:00
1600–1629	11:30	11:00	10:30	10:00	9:30	9:00	9:00	9:00	9:00
1630–1659	11:15	10:45	10:15	9:45	9:15	9:00	9:00	9:00	9:00
1700–0459	11:00	10:30	10:00	9:30	9:00	9:00	9:00	9:00	9:00
0500–0514	12:00	11:30	11:00	10:30	10:00	9:30	9:00	9:00	9:00
0515–0529	12:15	11:45	11:15	10:45	10:15	9:45	9:15	9:00	9:00
0530–0544	12:30	12:00	11:30	11:00	10:30	10:00	9:30	9:00	9:00
0545–0559	12:45	12:15	11:45	11:15	10:45	10:15	9:45	9:15	9:00

As can be seen from Table 2, the regulations consider both the number of sectors and the starting time of duty when limiting the FDP lengths. In addition, if crew is considered to be in an unknown state of acclimatization, maximum daily FDP should be as in Table 3, while when company utilizes FRM, the maximum daily FDP is allowed to be as in Table 4.

Table 3. Maximum daily FDP, crew members in an unknown state of acclimatization

Maximum daily FDP according to sectors						
1–2	3	4	5	6	7	8
11:00	10:30	10:00	9:30	9:00	9:00	9:00

Table 4. Maximum daily FDP, crew members in unknown state of acclimatization under FRM.

Maximum daily FDP according to sectors						
1–2	3	4	5	6	7	8
12:00	11:30	11:00	10:30	10:00	9:30	9:00

In addition, maximum daily FDP for acclimatized crew members with the use of extensions can be extended up to 1 hour, considering additional limitations.

The previous tables give example how EASA FTL rules consider flight duty structure when limiting maximum duty lengths. Depending on the supposed fatigue state of the crew member, different limits are used. These limits can be modified depending on the risk management processes of the company in question.

Rest is well covered in the EASA regulations. For rest periods minimum rest at home is defined as being at least as long as preceding duty period, or 12 hours, whichever is greater. This number is reduced to 10 hours when away from home base. Reduced rest is allowed when subsequent duties are adjusted accordingly. For time zone changes rest periods are adjusted accordingly, with additional local nights of rest at home base increased with increasing effect of fatigue due to time zone crossing and rest periods away from home base.

Extensions for flight duty periods can be granted by adjusting rest periods accordingly. With in-flight rest, the extensions can be granted in accordance with certification specifications applicable to the type of operation, considering sectors flown, in-flight rest facilities and augmentation of the basic flight crew. If unforeseen circumstances are met, commander's discretion is followed, which makes it possible to react to situations that would cause problems to operations in a larger scale.

The previous presentation of EASA FTL rules is given in order to emphasize the scope of rules and regulations used to calculate flight duty lengths and rest periods. In general the following methods and attributes have effect on the topic: starting time and place of current duty, starting time and place of next duty, time zone changes, number of sectors flown, length and location of previous and following rest periods, in-flight rest capabilities, augmentation crew available, break opportunities and FRMS. With all these issues affecting the duty and rest times the topic is complex to follow. Computerized schedule planning allows more complex rules to be created and monitored, which does bring flexibility to the operations, but it is difficult to understand if one is not specialized on the subject.

2.4. Planning of Crew Work Schedules

2.4.1. Principle

A crew work schedule is called a roster. It is built from work duties, which can vary from flight duties to different ground duties, for example simulator training or medical check. A roster is usually published a couple of weeks in advance. It can be planned for specific crew member, considering individual needs and requests, or a general roster can be planned that is then subjected to bidding system. There a crew member can bid for a specific roster according to a seniority list. Senior crew members get to choose first, while younger ones can choose last. In Europe the individual roster is favored over the bidding roster. (Barnhart et al. 2009)

A pairing is a duty that starts from a base of operations, usually an airport, and ends there. It can last several days, including nights away from home base. A flight leg, or a sector, is a flight from one airport to another. A pairing contains several flight legs that end the duty to the same base where it started from. Pairings form up a roster.

Planning can be divided into long term planning and short term planning. Long term planning is managed by crew planners, who create rosters for weeks in advance, which are published to the crew. Changes can be made after this, but are generally avoided. Short term planning is controlled by an operations centre, where the staff is responsible for the continuation of operations. The purpose is to ensure that every flight has a crew in changing situations.

2.4.2. Challenges

People may become sick or are unable to work due to exceeded duty times, which cause shortages in flight crew complements. The operations centre ensures that new crew members are located to fill the gaps and with as few disruptions to the scheduled operations as possible. Both short term and long term planning must fulfill the same rules in FTL. The difference comes from the timing and cost. Long term planning can make changes more freely and the cost of changes to the flight crew rosters is usually small. Changes made by the operations centre often incur overtime expense and may be hard to implement without disturbing operations and punctuality. Airline operators use stand-by systems in order to have reserve workers to call to work in case of sick leaves or other shortages of crew. As there are multiple rules to consider regarding duty and rest times, it may be difficult to find a flight crew member to fill in missing roles if the flight schedule is tight and work force fully utilized.

When rosters are planned, there is time to prepare them so that crew fatigue is considered. Pairings that would be too hard for crew are discarded and new ones are created to ensure alertness of crew and safety of operations. When people are called to work from stand-by duties or free days, there may not be enough time or resources to find a crew member that is well rested and will stay well rested for future operations. FTL rules are followed and illegal duties are not created, but there is a danger of increased fatigue due to the requirement of finding replacement crew in short notice.

3 Theory of Fatigue Models

3.1. General

Fatigue is generally modeled with bio-mathematical models that attempt to calculate the alertness level of an individual by taking into account the work hours, rest hours and circadian timing to estimate the fatigue levels of workers. It is not practical to build the models for individual preferences, so the models represent the average fatigue score of a group of people used to calibrate them. Every person has his or hers own personal needs for rest and recovery and people do not experience fatigue in the same way. While one person may be extremely tired after 6 hours of sleep, another person could manage well with that amount. Therefore fatigue models only give estimates of average person, and cannot be viewed as a simple truth.

The models in use today are used to compare rosters, evaluate new routes and evaluate changes in operational level. They can be used in accident investigations to better understand possible causes for incidents. What they are not supposed to be used for are firm go/no-go decisions, because of the nature of the models. They are only predictions and cannot calculate the true alertness of a person. In real life, every person is an individual, so generalized model cannot be used to make strict decisions based on individual characteristics and single fatigue score. In the future, the aim is to include fatigue risk models to the roster optimization phase, in order to build work schedules that are also optimized in the fatigue level. The earlier in the planning phase fatigue is considered, the more cost efficient solutions will be achieved. As the actual date of operations comes closer, any modifications to rosters incur costs and disturb already optimized schedule.

3.2. Anatomy of Sleep

Most of the fatigue models in use in aviation industry are based on the model of sleep regulation by Borbely (Borbely, 1982). This model was intended to explain the timing and duration of sleep as an interaction between two processes, sleep (process S) and circadian clock (process C), and is referred to as the Two-Process Model. Åkerstedt et al. (1997, 2004) added a third component of sleep inertia (process W) to further refine the model, which is generally referred to as the Three-Process Model.

3.2.1. Homeostatic pressure

Process S, also called homeostatic pressure, is the rising and falling of slow wave sleep, or deep sleep. This kind of sleep is essential for brain to handle memories and is linked to learning capabilities. A sleep occurs when the S reaches a high threshold, and wake-up occurs when the S drops below some low threshold. During sleep the S decreases in exponential fashion, and the pressure for slow wave sleep builds up during waking period. The longer one stays awake, the longer the length of slow wave sleep that will be needed in the next sleeping period. Across a sleep period the time spent in slow wave sleep decreases. During Rapid Eye Movement sleep (REM sleep) the brain activity is similar as during waking, and dreams are experienced during this time. Whereas the purpose of deep wave sleep is restorative, the purpose of REM sleep is still unclear. The process S has been proposed to be either exponential (Åkerstedt et al. 1997), linear with circadian variation (Hursh et al. 2004) or Gaussian (Jewett et al. 1999). (ICAO, 2011)

3.2.2. Circadian clock

Process C is a sinusoidal function that programs sleep to occur during night time and to stop during day time. The duration of this is approximately 24 hours and is called circadian rhythm, or circadian body clock. It is an internal period in human body, which is influenced by external factors, "zeitgebers", also called time givers, such as the cycle of daylight in local environment. In aviation this process causes problems when crossing time zones. The internal clock of a crew member is the same as at home, but the local time might be 12 hours in advance, meaning that when it is night at home and the body requires sleep, the sun is shining in the current location. Circadian rhythm can be measured by monitoring the core body temperature, which fluctuates by about 1ºC across the day. The daily minimum core body temperature corresponds to the time when people generally feel most sleepy. (ICAO, 2011)

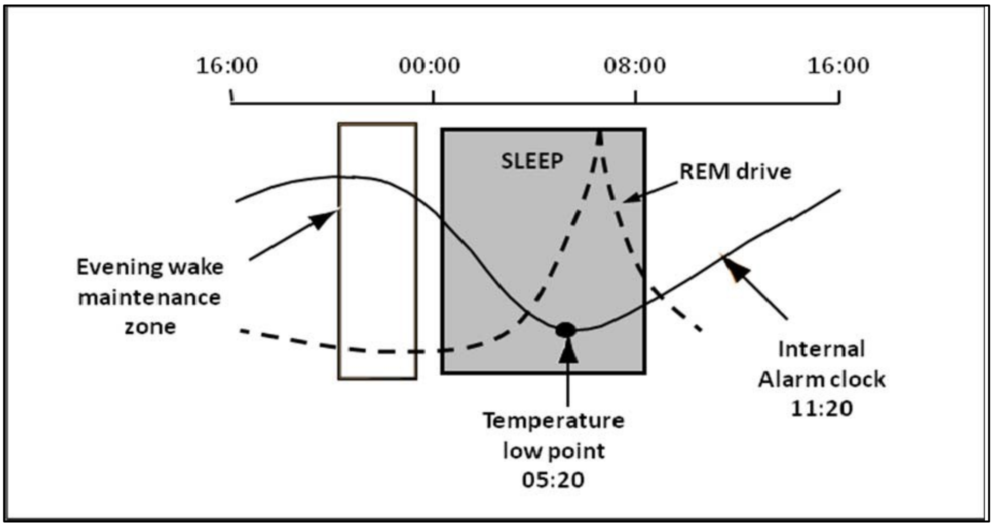

Figure 1. The circadian body clock and sleep after night duty (ICAO, 2011).

Figure 1 illustrates the circadian body clock throughout a 24-hour day. It summarizes the following features of sleep at night (ICAO, 2011):

1. Sleep normally begins about 5 hours before the minimum in core body temperature.
2. Wake-up normally occurs about 3 hours after the minimum in core body temperature.

3. REM sleep is entered fastest, and REM periods are longest and most intense, just after the minimum in core body temperature. This is sometimes described as the peak of the circadian rhythm in REM propensity (the dashed curve in Figure 1).
4. A variety of laboratory protocols have demonstrated that people are extremely unlikely to fall asleep 6-8 hours before the minimum in core body temperature. This has become known as the evening wake maintenance zone.
5. Laboratory studies also show that as body temperature begins to rise, there is an increasing pressure to wake up. This peaks about 6 hours after the circadian temperature minimum. This is sometimes referred to as an internal alarm clock, because it is very hard to fall asleep or stay asleep during this part of the circadian body clock cycle.

3.2.3. Combined daily cycle

The interaction between homeostatic process and circadian rhythm results in two times of peak sleepiness in 24 hours:

1. Window of Circadian Low (WOCL), which is around 3:00-5:00 for most people.
2. A peak in the early afternoon, around 15:00-17:00 for most people.

Restricted or disturbed sleep makes it harder to stay awake during the next afternoon sleep window. The precise timing of these windows of sleepiness varies between people who are morning types and people who are evening types. The morning types have the preferred sleep times earlier than the evening types, which results in differences in fatigue feelings if a work shift should start at the same time. As people grow older they tend to shift towards morning type behavior, which has been documented in flight crew members as well (ICAO, 2011).

The combined effects of homeostatic pressure and circadian biological clock result in windows when sleep is promoted (WOCL and afternoon peak) and windows when sleep is opposed (internal alarm clock in the morning and evening wake maintenance zone). Depending on the operations, these windows should be considered in rostering as they may affect flight crew rest possibilities. A morning duty requires person to fall asleep during evening wake maintenance zone, which might not be easy for most crew members. This leads to inadequate sleep, which results in greater fatigue levels.

3.2.4. Sleep inertia

Process W, sleep inertia, means the temporary disorientation and performance impairment after waking up from deep sleep. The severity is highest immediately after waking up, and the effects gradually decrease as time goes by. It can last as long as two hours in worst cases. This process is important when considering in-flight rests in long-haul operations, where the pilots are allowed to take short sleeps during flight duty. The rest opportunities increase overall alertness, but immediately after waking up alertness is low, which may increase operational risk, depending of the situation. During approach risk for incidents is on high level, whereas during cruise flight it is lower, therefore increased fatigue risk during cruise flight does not increase overall risk as much as during approach. In short-haul operations sleep inertia is not as influential as in long-haul, because in-flight rests are usually not utilized or allowed. Time from waking up in home or hotel to the start of work in cockpit is usually so long, that the effects of sleep inertia have normally worn off. (ICAO, 2011)

3.3. Structure of Bio-Mathematical Models

3.3.1. Inputs

Bio-mathematical models can be divided into two categories, one-step and two-step models, based on the input variables (Kandelaars et al. 2005). One-step models use actual timing of sleep and wake to predict fatigue. Two-step models use the work schedule as input and derive the sleep/wake pattern from that data. In the first step, the input work pattern is used to predict a probabilistic sleep pattern and a sleep/wake pattern is built. In the second step the estimated sleep/wake pattern is used to predict fatigue, as in the one-step model. Figure 2 illustrates the difference between the models.

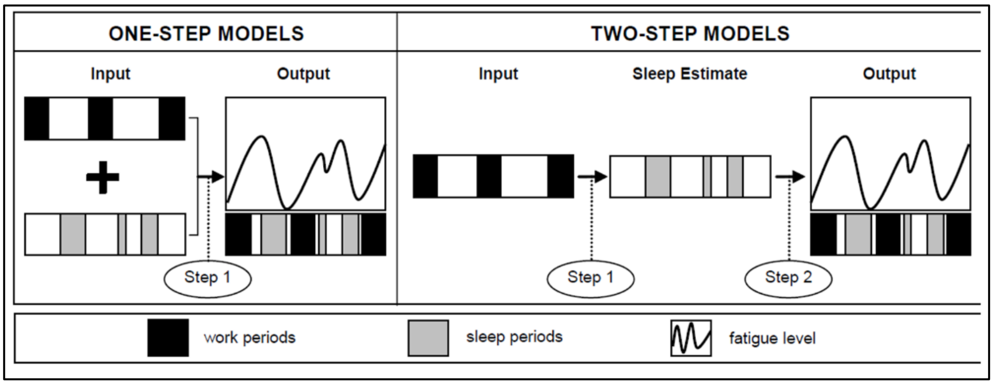

Figure 2. Comparison of one- and two-step models (Kandelaars et al. 2005).

Considering the characteristics of the one- and two-step models, it can be noted that the one-step model can be used only to estimate fatigue from past events, because it requires knowledge of the actual sleep/wake pattern. It gives more accurate results, but it is more difficult to use as the sleep data is hard to obtain. The two-step models are less accurate, because they build a probable sleep/wake pattern, but easier to use, because they use work schedules as the only input. Predicting fatigue on a future roster requires estimating the sleep/wake pattern, or using the work/non-work schedule, therefore models used to compare fatigue between future rosters have to be formulated as two-step models. (Roach et al. 2004)

A limitation in most of the current bio-mathematical models is that they predict fatigue for an average person. This does not correctly consider differences in individuals and their lifestyles. Any activities outside individual's roster are not reflected by the calculated fatigue score as they are not used as inputs in the model. In addition any task related differences could also be considered as inputs (etc. number of sectors in a duty, workload, and scheduled breaks) as those also affect the actual fatigue perceived. The more inputs are used, the more accurate the model can become. However, it also becomes harder to implement, because not enough data may be available.

3.3.2. Outputs

Most of the bio-mathematical models provide a fatigue or an alertness prediction value over a given work period (Branford et al. 2014). The Karolinska Sleepiness Scale (KSS) is a scale ranging from 1 ("very alert") to 9 ("very sleepy, great effort to keep awake"). The Samn-Perelli (SP) fatigue scale is a 7-point scale with scores ranging from 1 ("fully alert, wide awake") to 7 ("completely exhausted, unable to function effectively") (Samn & Perelli, 1982). These both give subjective scores for sleepiness, but have been evaluated with objective measurements (Kaida et al. 2006) or used in studies (Powell et al. 2007).

The advantages of the KSS and SP are the ease of use in operations and transparency to crew and planning. A subjective and well defined score is easy to collect from crew and they understand it well. However these two meters are not interchangeable, because the SP extreme refers to extreme fatigue while KSS extreme is a very low level of alertness. Another issue is related to risk. A specific score of fatigue cannot produce a specific level of risk for a task without considering the demand associated with the task (Branford et al. 2014).

There has been research to develop metrics more relevant to the real world (Dean et al. 2007). Rangan et al. (2013) propose a method were the fatigue risk is proportional to the time spent below a fatigue threshold. Therefore a long time under a given fatigue score threshold increases the risk, independent of the specific task at hand. Cognitive effectiveness is proposed by Hursh et al. (2004) to measure effects of fatigue in operational activities. It is interpreted as an inverse of fatigue and ranges in score from 0 to 100. This metric is derived from Psychomotor Vigilance Task (PVT), a reaction time test used to measure neurobehavioral performance. Another approach is developed by Dawson and Fletcher (2001) where a fatigue score is calculated for a time window, on a scale from 0 to 140. Working Monday to Friday 09:00 to 17:00 is represented by a score of 40, set by validation studies. If a work schedule gives score below 80 it is considered to be acceptable, while a score of over 100 is considered to be unacceptable, requiring countermeasures against fatigue. Whatever the fatigue risk score is, it should consider more than just KSS or SP scores, which contain only little information regarding risk analysis.

3.4. Fatigue Models in Use

In this section, several current models are briefly compared. The purpose and scope of these models vary largely, but all of them have been validated for use in aviation environment. BAM, CAS and FAID have been created for commercial use and have the highest variety of qualities and support. FRI, SAFE and SAFTE-FAST have been originally commissioned by government authorities. SWP is the simplest of the models while FRI and SWP are free for download.

Boeing Alertness Model (BAM)

BAM is a bio-mathematical model built on the three-process model of alertness augmented with advanced sleep prediction. The output is based on KSS, which is converted to an alertness score on a scale from 0 to 10,000. Four large scale data collections have been undertaken to validate the model, and with data shared by airlines, close to 60,000 assessments from actual operations have been used in refining the model. It is possible to integrate BAM with crew planning software. It can also be used via mobile applications to gather data and monitor individual fatigue. (Jeppesen, 2009; Branford et al. 2014)

Circadian Alertness Simulator (CAS)

CAS estimates fatigue risk based on an individual's sleep-wake-work pattern. It is built on the two-process model of alertness and can use actual sleep history as an input or simulate the sleep based on work patterns. The output is a Fatigue Risk Index between 0 and 100. Model has been validated in various transportation fields, including railroad, trucking and maritime, and optimized for aviation specific use. It is targeted for crew planning applications. (Moore et al. 2004)

Fatigue Assessment Tool by Interdynamics (FAID)

FAID uses working hours as an input while the output is a fatigue score on a scale from 0 to 140, indicating different levels of fatigue exposure for different working hours. A higher score means higher fatigue exposure and provides an indication of the likelihood of performance impairment associated with fatigue. Fatigue tolerance levels are used to limit the working hours based on predetermined levels. Data from Australian train drivers was used to develop the model. (InterDynamics, 2014)

Fatigue Risk Index (FRI)

Used for comparing work schedules and identify the fatigue risk of a shift. The outputs are a fatigue index and a risk index. The fatigue index is based on KSS, multiplied by 100, which describes the average probability of value seven in the KSS. The risk index is an estimate of the relative risk of making an error that could contribute to an accident. The model was validated with data obtained from aircrew, train divers and industrial shift workers. (HSE, 2006)

System for Aircrew Fatigue Evaluation (SAFE)

Built especially for aircrew fatigue evaluation and based on the two-process alertness model. Gives a SP fatigue score generated for every 15 minute interval in a flight duty schedule and predicted likely sleep patterns. Data used to validate the model was collected from pilots working different schedules with different airlines. The model can be used in conjunction with crew scheduling optimizers. (CAAUK, 2007)

SAFTE-FAST

Acronym for sleep, activity, fatigue and task effectiveness (SAFTE) and fatigue avoidance scheduling tool (FAST). Model provides several performance metrics (e.g. percent change in cognitive speed, lapse likelihood, reaction time) and sleep-wake metrics (e.g. sleep reservoir, circadian phase), with outputs measuring duty time and critical time below adjustable fatigue risk criterion line. Model has been validated with railroad and aviation workers and with people under laboratory settings. (Branford et al. 2014)

Sleep/Wake Predictor (SWP)

This is a bio-mathematical model based on the three-process model of alertness. Output is a predicted alertness curve, in a 1-21 point generic scale or KSS. Total time of work above critical limit is calculated and used to summarize the risk of a particular work schedule. Inputs consist of work patterns or sleep/wake patterns.

The SWP has been validated in a number of studies while the underlying three-process model has been validated against EEG parameters and under laboratory performance tests. The model is suitable for assisting schedulers to evaluate fatigue, but not for large scale roster development. (Åkerstedt et al. 2004)

4 Model

4.1. Empirical Data

4.1.1. Data gathering

A survey for a pilot test group was conducted to gather data for the evaluation and analysis of the fatigue model. The test group consisted of 12 pilots, including both captains and first officers. They were briefed on the goals of this project and the importance of good results. Emphasis was given to make sure that the test group understood that the results were only to be used in this thesis, with no information distributed to the company or follow-up conducted regarding the results of the fatigue survey. Best results are achieved when the test group understands that regardless of their answers, no measures are taken even if they make mistakes in their duties. This is also general principle in aviation industry, where goal is to improve operations and not punish those who report their actions.

The survey was filled on an iPad application which is also used for reporting other issues and filling in forms necessary to the flight operations administrational functions. The main concern was to make the fatigue form quick and easy to use, so that the pilots would feel it was not taking too much time from them to fill it out. The iPad application was suitable in this regard, because it was always available and familiar.

The period for the survey was one month during summer. It is peak season, which is always hectic, and requires lots of effort from flight crew. Rosters are hard and vigilance may be compromised if the crew does not get enough rest. It is a good moment to measure fatigue as the results should show the effects more clearly. Due to summer vacations, the survey could not be published for every participant at the same time. The length of the survey period was one month for everyone, but the periods were placed throughout summer.

The pilots were required to give a fatigue score before and after duty, with the possibility to give scores also during duty. As the operations were short-haul, the fatigue effects should be quite linear between the start and end of duty. Because the time zone changes are not as dramatic as in long-haul operations, crews internal clock is tuned to the crew base time zone and diurnal variations are minimal. In this regard, the fatigue scores estimated during duty are not used to assess fatigue in the developed model. In long-haul operations the length of flights and duties is much longer, time zone changes are more frequent and available day light can be minimal due to local time. The operations are very different and require different modeling parameters.

4.1.2. Survey form

The main structure of the fatigue survey form, and explanations why each attribute was chosen, is listed below. Figure 3 shows a screen capture of the survey, completed on an iPad. The form extends further by scrolling the screen.

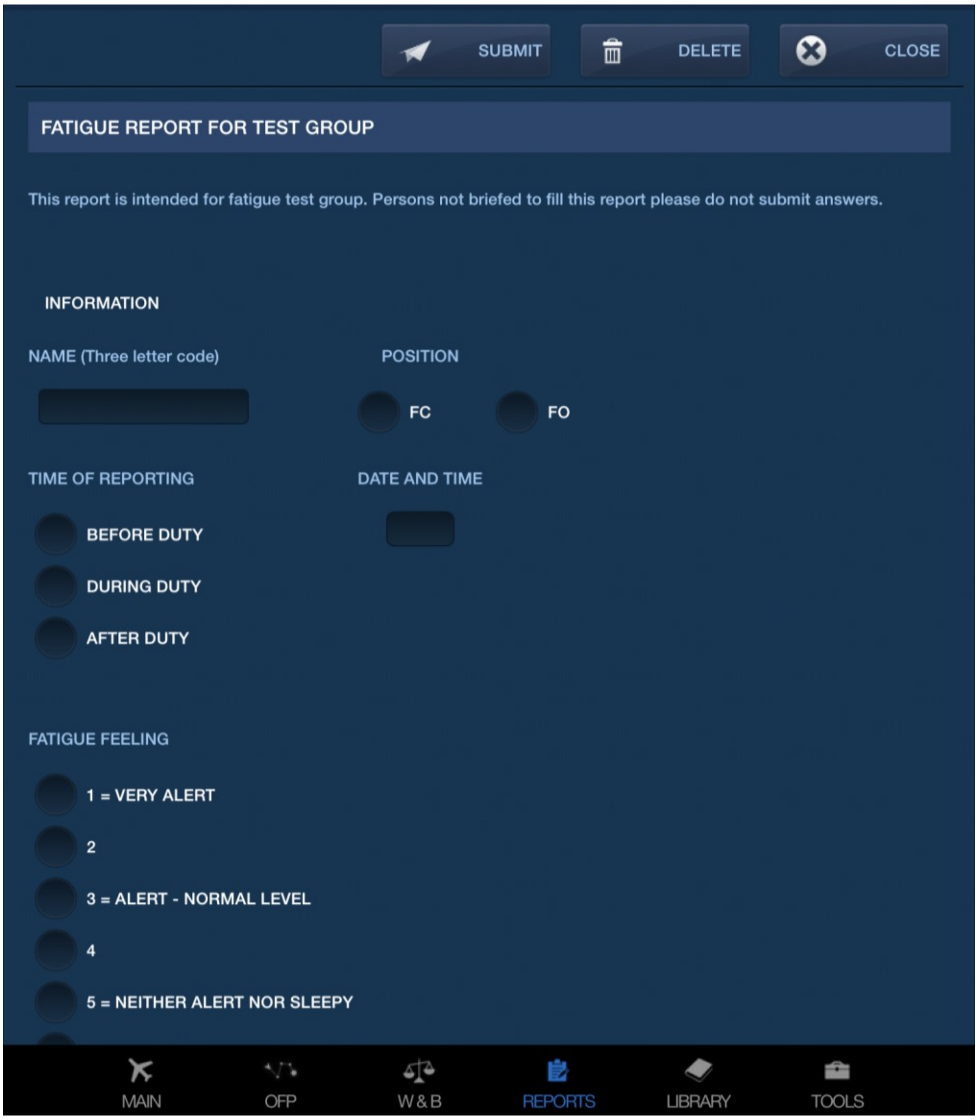

Figure 3. Screen capture of the fatigue survey form.

Time of reporting

This question asks whether the fatigue score was given before, during or after duty. When comparing the before and after scores, a drop in the fatigue score can be calculated and used for analysis. This data also gives the time of reporting, whether it is early morning or late afternoon, which should have large effect on the fatigue score.

Fatigue feeling

This score is based on Karolinska Sleepiness Scale, which measures subjective fatigue of the crew member on a scale from 1 to 9.

Amount of sleep on previous night

This question gives data on the amount of rest achieved before duty. Although this kind of data is not available for roster planners, because it is not known beforehand how long crew members sleep, it can be used to analyze results.

Possible other than work related issues

This is a yes or no question, which gives indication if pilot feels that there might be something causing lower than normal fatigue, but which is related to his personal life. This information can be used to analyze deviating results.

Free text

A text field is given to write down possible reasons that could have had effect on fatigue, such as weather conditions, difficult destinations, waiting time or large number of legs.

Outside of the survey was left the structure of the roster. This information is gathered from the rostering system in order to keep the survey as short as possible.

4.2. Model Formalization

4.2.1. Sources of fatigue as inputs

As the model in question is supposed to predict fatigue that is caused by the crew rosters, the only input available is the work schedule, meaning duty starting, duration and ending times. From the input it is possible to calculate duty length, cumulative duty time, rest time, cumulative rest time, time of day and number of sectors flown. These six attributes should have the most effect on the alertness state of crew members based on previous research and legislation. Figure 4 presents an example of the input signal of the model, where x_w is 1 when person is working and 0 otherwise.

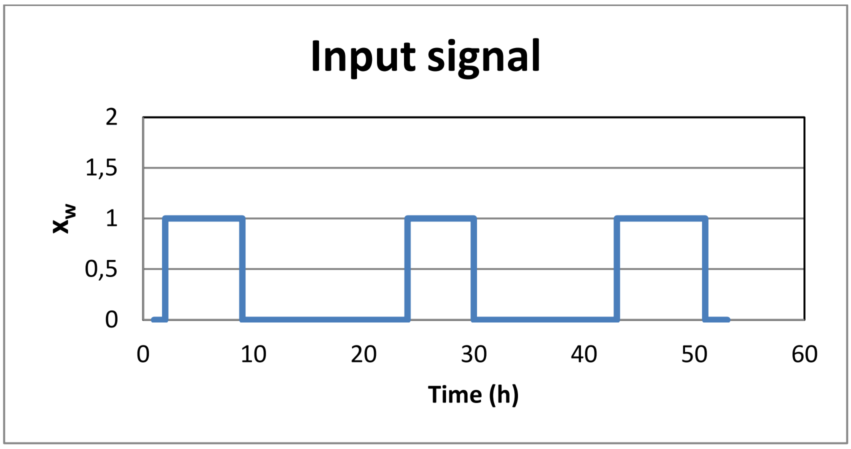

Figure 4. Example of an input signal of duty hours.

Duty duration times have impact in the sense that the longer a person is working and concentrating hard, the more tired he or she will feel in the end. Cumulative duty time measures the cumulative amount of duty hours. In FTL this must be kept below safe levels in order to protect against fatigue. The effect is more hidden, as the impact will build up. After a long work day a person will feel tired regardless of the previous days. But if there are several long days following each other, a person will most likely feel tired sooner than if the previous days had been free days.

Rest times have a direct impact on fatigue levels. If there is not enough rest allowed to the crew, they will feel exhausted soon. One short rest may be easy to handle if there is a longer rest period following, but if the cumulative rest times do not allow enough time to recover from work, then crew fatigue levels will begin to rise.

Time of day is connected to the circadian rhythm, which affects how a person is able to sleep and how alert he or she feels. If a duty starts during WOCL, a person will most likely feel tired, even though there is an adequate sleep achieved. Same happens if a duty ends during WOCL, as then there is a risk for high levels of fatigue due to natural rhythm of body.

The number of sectors flown is directly linked to the amount of work load. Most of the work is done during take-off and approach, while steady level flight is mostly monitoring. Considering a same duty period with different number of sectors flown, the one with fewer sectors is considered to be easier and less tiring.

Long haul crews face a different issue with long periods of constant monitoring which is very difficult for humans. With short haul, the number of sectors is more clearly linked to the intensity of workload and perceived fatigue. Table 2 in chapter 2 shows that the current FTL considers the impact of WOCL and number of sectors to the crew alertness by limiting the maximum flight duty period depending on the starting time and sectors flown. In that regard it is clear that a bio-mathematical model should possibly consider them as well.

4.2.2. Output of the model

Output of the model is defined as an alertness score, AS, ranging from 10 to 90, 90 meaning fully alert and 10 being completely exhausted. The KSS scale is fitted into this range by inverting the scale. This transformation is done in order to create a scale that is easier to interpret. In the alertness scale a fatigue score of 1 in KSS is equal to 90 and 9 in KSS is equal to 10 in alertness score. The transformation is defined as

$$AS = 10(10 - KSS) \tag{1}$$

The alertness score is calculated as a sum of the homeostatic process S and circadian process C, as shown in Figure 5. Process S’ is the recovery phase from S, where the homeostatic pressure relieves and recovery occurs. All three processes give individual alertness scores that are summed to calculate the final alertness score, which is the output of the model.

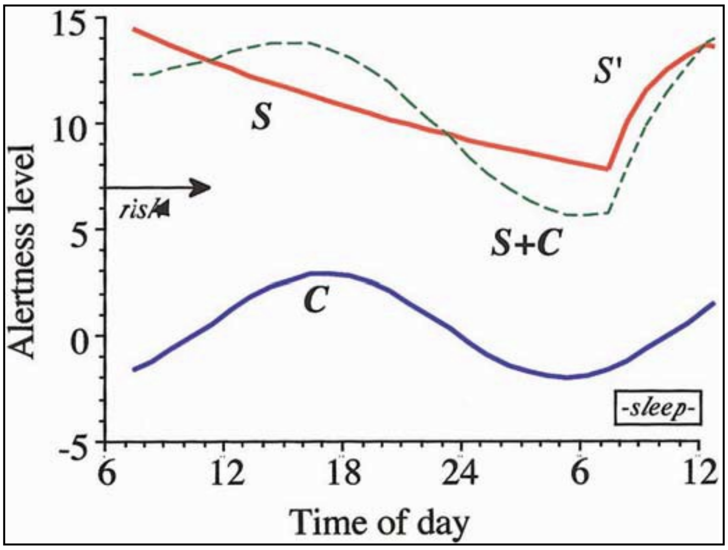

Figure 5.The components of the three process model. (Åkerstedt et al. 2004)

4.2.3. Modeling circadian rhythm

Process C, the circadian rhythm, is modeled as a sinusoidal function,

$$C(t) = A\sin(Bt + \varphi), \tag{2}$$

where t is the time of day, $t \in [0,24[$, $A \geq 0$ is the amplitude of the wave, B is the frequency and φ is the phase. The minimum of the process C is set at 5:00 in the morning and with the frequency being 24 hours, constants are calculated to be $B = \frac{\pi}{12}$ and $\varphi = \frac{13\pi}{12}$. With the minimum being at 5 am, the maximum of the wave is at 17:00 in the afternoon. This fits well with the evening maintenance zone, where it might be difficult to fall asleep.

The effects of waking up early are incorporated into the process C, because the alertness score of C is low in the morning and rises as the day goes further. In similar effect duties that end very late get low alertness score from the process C because the score starts to fall after late afternoon and shifts ending after midnight get similarly low scores as early morning shifts.

Amplitude A defines the weight of the process C compared to process S. There must be a balance between these two processes, because both have impact on the alertness levels and are required to calculate the final score. The value of the constant is defined later.

4.2.4. Modeling homeostatic process

The homeostatic process S is formulated as a linear function. If a person is working, fatigue rises in a linear fashion. The recovery process S' is formulated in an asymptotic fashion. When a person is not on duty his or hers alertness increases, but there is a limit to how alert a person can be. This is handled so that when fatigue lowers, the speed of recovery reduces. A fully alert person will not gain alertness through rest. If the model would have a component for actual sleep process, the recovery should be formulated in exponential fashion, as it has been studied that the recovery is faster during the early hours of sleep (Åkerstedt et al. 2004). Both S and S' depend on the length of the period but not on the circadian timing. Work place studies have shown that the longer the rest period, the more sleep people are able to collect, and the longer the duty period, the more they experience fatigue (Roach et al. 2004). In the homeostatic process S a cumulative component is added to include the effects of previous days. In recovery process S' this effect is not included, but the recovery is assumed to be dependant of available rest time and current predicted fatigue level.

Processes S and S' are defined so that

$$S(x_{w,t}) = Dx_{w,t} \tag{3}$$

$$S'(x_{w,t}) = F \cdot E \cdot (1 - x_{w,t}), \tag{4}$$

and the total homeostatic pressure at time t is defined as

$$H_t = S_t + S_t' + H_{t-1} \tag{5}$$

$$= S_t + S_t' + S_{t-1} + S_{t-1}'. \tag{6}$$

Coefficients D and F are formulated so that

$$D(x_{wcum}, x_{rcum}) = D_0 + G(x_{wcum}, x_{rcum}) \tag{7}$$

$$F(H_t) = \left(1 - \frac{H_{t-1}}{ul}\right), \tag{8}$$

The notation is as follows: $x_{w,t}$ is 1 if crew member is working at time t and 0 if not, x_{wcum} is the cumulative duty hours and x_{rcum} is the cumulative rest time in previous consecutive work days. D_0, G, and E are constants derived later. The coefficient $F(H_t)$ depends on the current fatigue score without process C, and ul

is the upper limit for the fatigue score. The cumulative duty hours are calculated from previous work days so that if rest time exceeds 24 hours, the x_{wcum} drops to zero. This assumes that one day of rest allows crew to recover themselves from cumulative effect. The cumulative rest hours are calculated in the same way, calculating the rest before duty and adding the values until a rest day is rostered.

4.2.5. Total alertness score

We can now define total alertness score at time t, with initial alertness score S_0, as

$$AS(t, x_{w,t}) = C(t) + H(x_{w,t}) \tag{9}$$

$$= C(t) + S(x_{w,t}) + S'(x_{w,t}) + S(x_{w,t-1}) + S'(x_{w,t-1}) \tag{10}$$

$$= C(t) + S_0 + \sum_{t=1}^{t} \left(S(x_{w,t}) + S'(x_{w,t})\right) \tag{11}$$

The result is a function consisting of sinusoidal, linear and regressive functions. The sinusoidal part is same for every day, though it is argued that the circadian rhythm will move if people work continuously morning or night shifts. In this thesis we will assume it to be independent of work history. Function S is linear when considering individual duties, where variables x_{wcum} and x_{rcum} have fixed values depending on the work schedule. Function S' is regressive as it depends of the previous fatigue score. Variable x_w is the input signal to the fatigue model, from which all other variables can be calculated.

When a crew member is working, the function S decreases the alertness score in the function H. When the duty ends, the alertness score begins to increase according to function S'. The sum of these functions over time is the effect of the homeostatic process. The circadian rhythm is independent of previous values; it depends only of the current time. Calculating the sum through time we get the current total alertness score.

4.2.6. Limiting alertness score

A few limits must be set to the alertness score function in order to keep it in the defined range. The score must be between the scale from 10 to 90, meaning that however long a crew member rests, he or she can not score above 90 or work so much that the score goes below 10. The asymptotic nature of recovery function limits the alertness score to ul but the function $H(t)$ must be refined further to include the lower limit ll. We shall define

$$H(x_{w,t})_{max} = \max\left\{ll, S_t + S'_t + H(x_{w,t-1})_{max}\right\}. \tag{12}$$

We can now write the final equation with limitations as

$$AS(t, x_{w,t}) = C(t) + H(x_{w,t})_{max}\,. \tag{13}$$

This formulation keeps the score from the homeostatic process inside a scale $[ll, ul]$. The sinusoidal function $C(t)$ increases or decreases the alertness score in a range $[-A, A]$ so the maximum and minimum for the alertness score is on a scale $[ll - A, ul + A]$. The lower limit is forced with the maximum operator in equation (12). Because the alertness score is assumed to reduce in linear fashion, there must be a lower limit which is not breached.

The circadian process is kept separate from the homeostatic process due to the modeling of circadian rhythm. In a situation where a person has had several rest days and is calculated to be recovered by homeostatic process, an early check in should not result in very high alertness score. The circadian process takes into account the early wake-ups and late check-outs, because even a fully rested person feels the effects of one's biological clock. As in Figure 5, the process C gets negative values, so that even when fully rested by the recovery process S', the total alertness score will not be 90 if duty starts very early.

4.3. Analysis of Data

4.3.1. Key figures

The data collected contains information from 12 pilots. There is data from 147 different flight duties, which include 133 fatigue scores from before duty, 116 fatigue scores from after duty, and 104 calculated differences between before and after duty fatigue scores. The missing fatigue scores from either before or after duty prevent calculating rest of the difference scores. The mean length of duty was approximately 8 hours, with the average before duty fatigue score being 3.8 and the after duty score 5.1 in KSS. The average check-in time was at 10:00 and check-out time at 18:00. The average number of legs was 3.3 per duty. Reported sleep was on average 2.5, which is quite high, meaning that pilots sleep less than 7 hours on average. The impact of personal life to sleep was on average 0.2, which is low, meaning that 2 out of 10 reports had some issues in personal life that they felt had an effect on fatigue.

The biggest challenge was to get consistent answers because the pilots tend to forget to fill in either before or after duty fatigue scores or both completely. Also some persons are more likely to give answers than others, so a greater share of data may come from some individuals. This becomes large problem if the acquired data is not good, meaning situations where an individual who is keen on answering does not fully understand the differences of the answer options. However, everyone experiences fatigue differently, so the problem of generating average model that suits everyone is in itself difficult.

4.3.2. Feedback from flight crew

The fatigue questionnaire for the flight crew also included a text field for commenting on reasons that might explain their current fatigue score, or for giving feedback on the subject. Several factors affecting fatigue were given, here are listed the most important ones:

1) The single most listed reason for increased fatigue was the change from evening shift to morning shift. A duty ending late in the evening followed by a morning duty contains also short rest between these duties. The short rest accompanied with early check-in results in higher than normal fatigue scores with flight crew.
2) Early morning shifts in general, possible increased effect if there are several in conjunction or previous day has been long. Early wake-ups require pilots to go to sleep earlier in the evening. However, falling asleep may not be easy then.
3) Long duties ending late in the evening.
4) Difficult rotation with shifts. Example given of a shift where the next duty starts on the next day at the same time as the previous ended, followed by

another similar cycle. This creates a 24-hour rest between duties, which is not easy to switch to.

Other reasons listed were long duties, too short sleep period, bad flight weather and general problems in operations that require extra effort to handle.

Because the purpose of the model is to predict fatigue based on crew rosters, some of these fatigue sources cannot be included in it. Sleep quality and amount is possible to model, but it requires information which is only available after each night or rest period. External factors, such as weather, could be included in the model, because weather forecasts are available and quite accurate. Predicting the fatigue for same day with up-to-date weather information is possible and could give more accuracy to the model. In the same regard, utilizing up-to-date sleep information from crew members could be used to calculate more accurate fatigue predictions for upcoming days. The purpose of this work is, however, to predict fatigue further into the future and help roster planning, which is carried out weeks before the actual date of duty. Therefore external, real time effects are left out of the model.

4.3.3. Effects of sleep

Figure 6 shows the effect that the amount of sleep on previous night has on the perceived fatigue score before duty on the next day. As the amount of sleep reduces, the fatigue score increases. The immediate effect of short or bad sleep is easily recognized from this figure. The problem for planners is that it cannot be defined beforehand how well the crew members use their rest time.

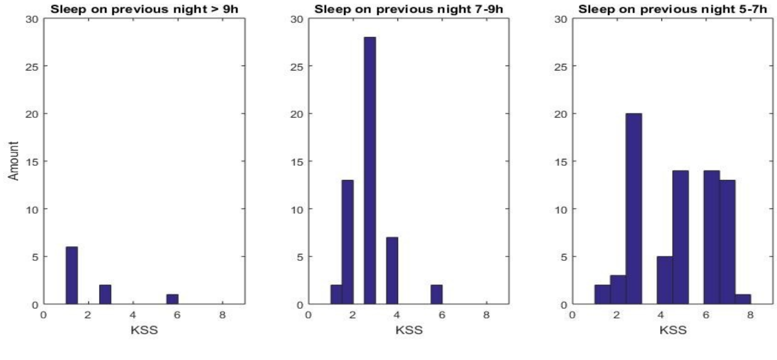

Figure 6. Distribution of before duty fatigue scores with different amounts of sleep.

Figure 7 presents the differences with different sleep amounts when comparing the fatigue score to the check-in time of the pilots. With more than 9h of sleep the fatigue score stays low, but as the sleep amount reduces so does the fatigue score increase with early check-ins. From the data it is not possible to say is the early check-in time reason for high fatigue and short sleep, or is the short sleep the main factor. It is clear that with early check-in times the fatigue scores are higher than with later starts of duty. With normal average amount of sleep between 7 to 9 hours, the fatigue score before duty is mainly concentrated in the score between 2 to 4, but there are few very early check-in times in that data group. Most of the early check-in times are present in the data group with sleep from 5 to 7 hours and the fatigue scores are higher in there. In Figure 8 the effect on fatigue score at the end of duty is similar. When the amount of sleep is reduced, the fatigue scores increase and higher scores are reported with earlier check-out times than with longer sleep times. It is clear that with reduced rest pilots feel more tired earlier than with adequate rest.

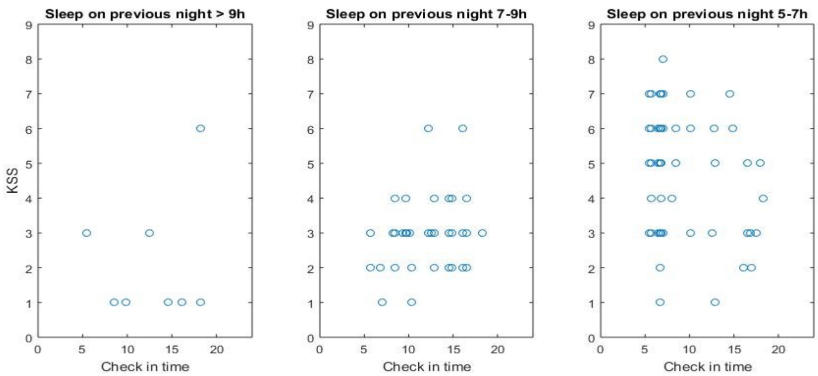

Figure 7. Effect of sleep amounts on check-in time fatigue score.

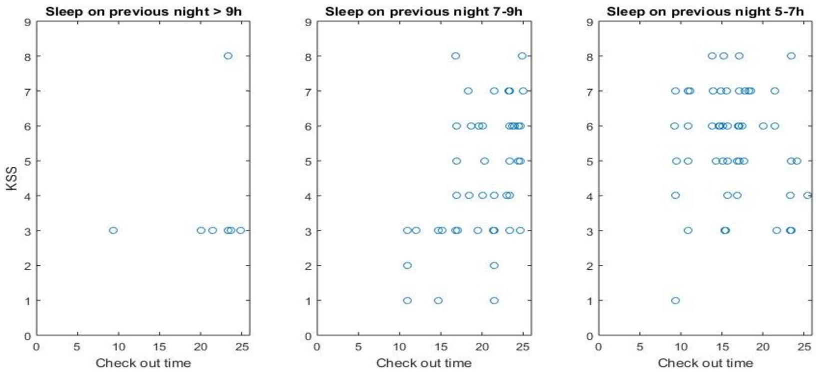

Figure 8. Effect of sleep amounts on check-out time fatigue score. Time 25 means 01:00 in the night.

4.3.4. Check-in and check-out time

The effect of check-in time is clear in Figure 9, where it is plotted average fatigue score versus check-in times in two hour time windows throughout a day. Early duties suffer from a higher fatigue score, but after 7:00 the average score flattens and is fairly constant for the rest of the day. For flight duties starting before 7:00 the mean fatigue score is close to 5, which means “neither alert nor sleepy”, whereas for the rest of the day the average is closer to 3, which is “alert – normal level”. This is in line with the assumption that regardless of work history, an early start raises the fatigue score. Another increase in fatigue is seen in duties starting after 17:00. In the same figure the average fatigue score for check-out times is also plotted. The results are more random and no clear result can be seen on the graph. The average score is more scattered throughout the day, and no conclusion can be made from this.

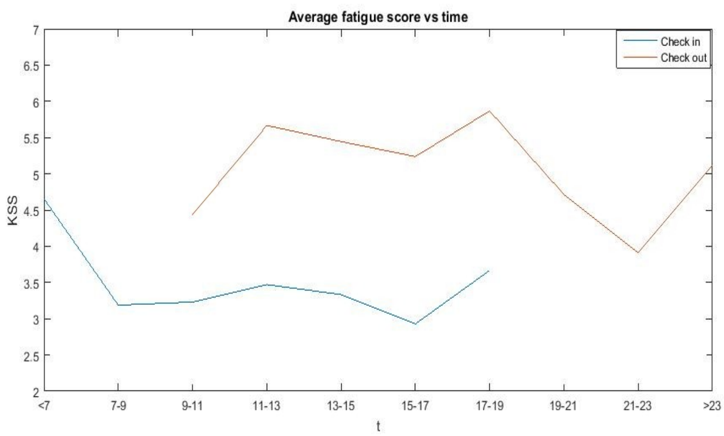

Figure 9. Average fatigue score versus check-in and check-out times.

4.3.5. Effects of duty length and previous rest

The effect of duty length should be straightforward, the longer a duty is the more crew alertness drops during it. In Figure 10 is plotted the difference between check-in and check-out time fatigue. A simple linear trend is dropping, but for many instances the difference in the fatigue scores between before and after duty is zero or even positive. Pilots often feel that in the end of a duty they have a similar level of alertness as in the beginning. The positive values can be explained due to the early check-in times. Early in the morning pilots feel very tired, but as the day progresses their alertness rises. With linear fitting it can be calculated that on average a one hour increase in duty length increases the fatigue score by 0.2 in Karolinska Sleepiness Scale.

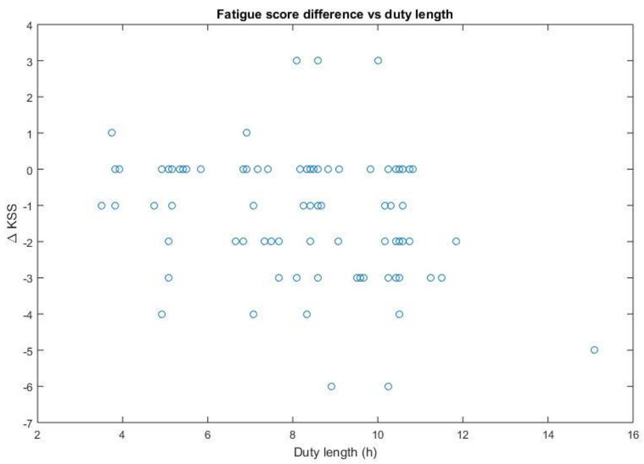

Figure 10. Effect of duty length to fatigue levels.

First graphs in Figure 11 and Figure 12 present the average fatigue difference between duty length and previous rest times. The other graphs are the check-out time fatigue versus duty length and the check-in time fatigue versus previous rest time. These pictures show that increasing duty length increases fatigue and the longer the duty the higher the drop in alertness. The effect of previous rest time is not clear, because the average check-in time fatigue is fairly constant regardless of the amount of previous rest. The fatigue difference plot shows some increasing trend, which would mean that increasing rest time reduces the drop of alertness during next duty.

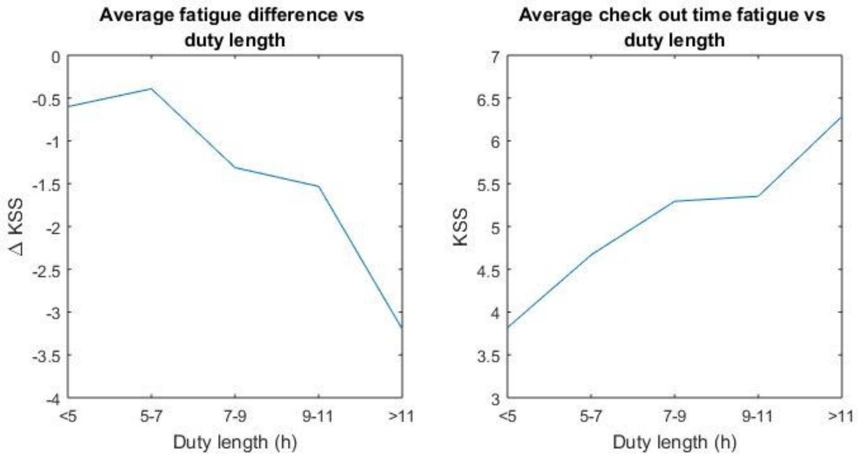

Figure 11. Average difference in fatigue score between check-in and check-out times versus duty length, and average check-out time fatigue versus duty length.

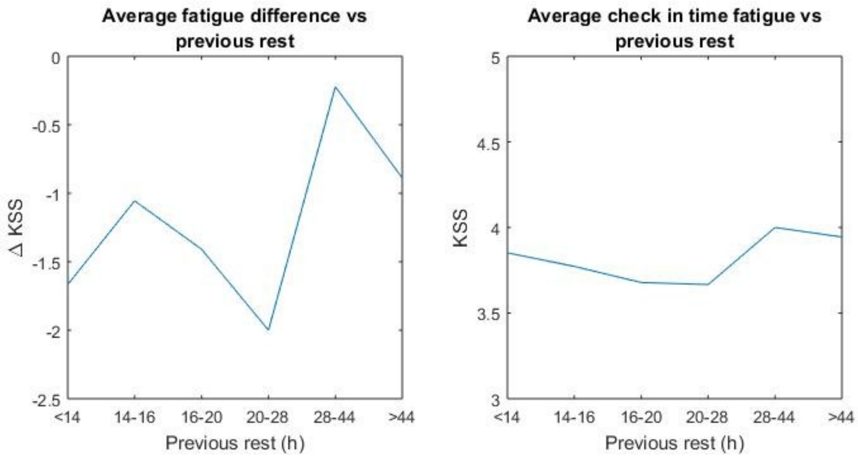

Figure 12. Average difference in fatigue score between check-in and check-out times versus previous rest time, and average check-in time fatigue versus previous rest time.

The effect of sleep on fatigue was clear from the data, but the amount of sleep is regulated by crew members themselves. Effect of rest times is not as clear, but roster planners can only plan the check-in and check-out times and rest times, and in the end it is up to the flight crew to get enough rest.

4.3.6. Cumulative effects

Cumulative duty hours and cumulative rest times are calculated from the shifts before current work shift. A consecutive work day adds to the cumulative scores and a rest day sets the scores to zero. Both variables are therefore calculated backwards until previous rest day. Figure 13 shows an example of a work schedule, where W_i marks a duty and R_i marks a rest. On the work day W_3 the cumulative duty hours would be $W_1 + W_2$ and the cumulative rest times similarly $R_1 + R_2$. On the work day W_4 the cumulative values are zero.

Free day	W_1	R_1	W_2	R_2	W_3	R_3	Free day	W_4	R_4

Figure 13. Roster example for calculating cumulative values.

Figure 14 shows the effect of cumulative duty hours and Figure 15 the effect of cumulative rest times against before and after duty fatigue levels. From these figures it is not possible to see any clear trend that the increase of cumulative work hours would increase fatigue or that the increase of cumulative rest time would decrease fatigue. With cumulative duty hours in both before and after duty graphs the fatigue scores do not shift notably to top right with increasing duty hours, as is expected. Same issue is with cumulative rest times, as the increase of rest times does not reduce fatigue. The scores do not move bottom right in either graph with increasing rest times. Due to this, no assumptions can be made from these graphs, other than that there is possibly no cumulative effect present in the data.

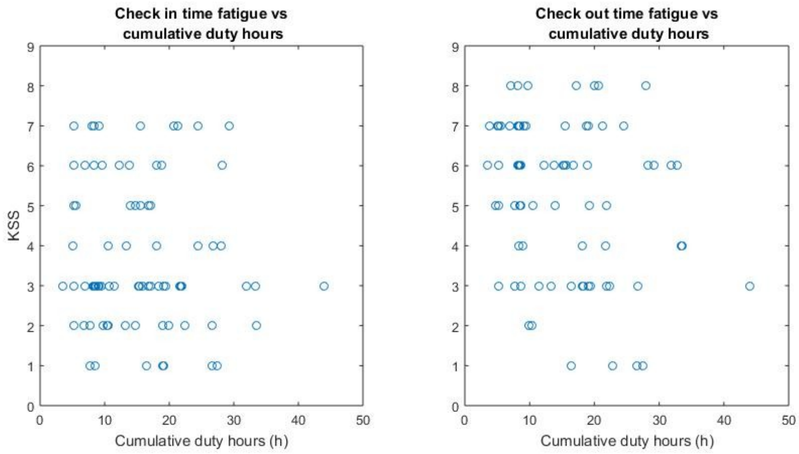

Figure 14. Effect of cumulative duty hours in consecutive work days to fatigue score, before and after duty.

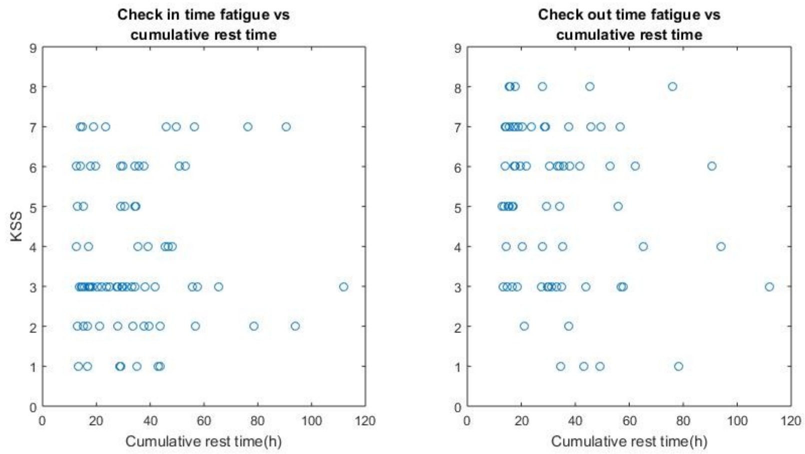

Figure 15. Effect of cumulative rest time in consecutive work days to fatigue score, before and after duty.

Comparing the average difference between before and after duty fatigue scores versus the cumulative duty hours and cumulative rest time should reveal the effect of cumulative factors. The first graph of Figure 16 presents the effect of increasing cumulative duty hours. The more work crew member has done, the lower the difference. This seems odd, because the increase of cumulative duty hours should increase the fatigue build up. In the second picture the effect of increasing cumulative rest time is more as expected, as there is some trend towards reduced difference with increasing rest times. Both the graphs are not very clear in results, but it is possible to make some interpretations. One possible explanation is that as the cumulative duty increases, the difference does not increase, because the crew gets used to working. Fatigue is felt stronger in the beginning of consecutive duty days than at the end, which results in lower difference values, but the overall alertness is lower.

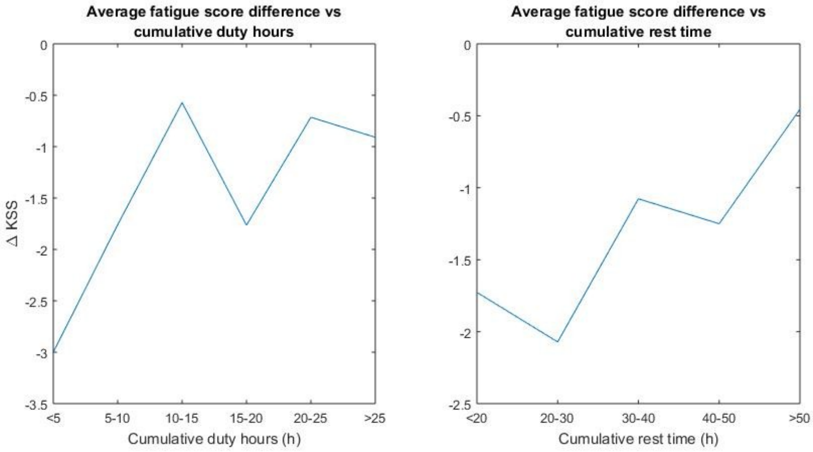

Figure 16. Effect of cumulative duty hours and cumulative rest time in consecutive working days to fatigue score difference.

Because of the randomness of the results, a variable R_{cum} is created, which is the ratio of cumulative duty hours per cumulative rest time for current duty, which are calculated as in example in Figure 13. Both before and after duty fatigue levels have therefore same R_{cum} value. This variable is defined in order to better understand the cumulative effect in fatigue levels. Figure 17 presents the R_{cum} value versus average fatigue difference. The results are better than with separate cumulative variables, because increasing ratio increases the fatigue difference between check-in and check-out times.

The ratio describes the amount of work versus rest, so with high values the person in question has long duty hours with little rest. A value close to one means that for every 8 hour duty, only 8 hours of rest is available, in general. For example, a 12 hour duty can be followed by a 12 hour rest according to FTL rules, but in long term it is not a good work schedule in terms of human factors.

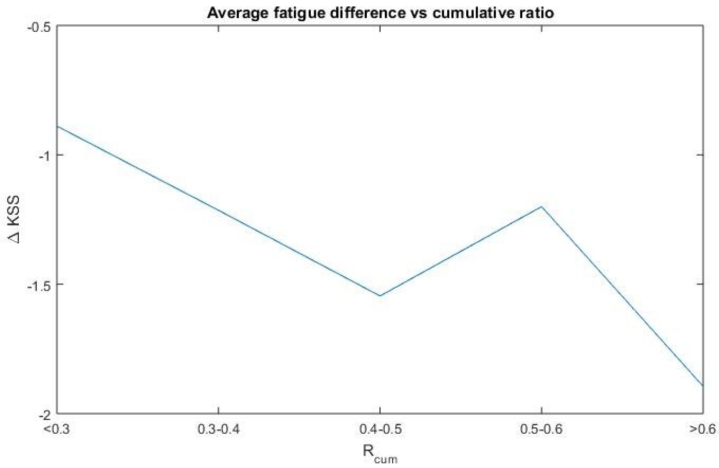

Figure 17. Average fatigue score difference versus ratio between cumulative duty hours and cumulative rest time.

4.3.7. Effect of the number of legs

Figure 18 presents the number of daily legs versus average after duty fatigue. The effect should be that increasing the number of legs should increase the fatigue levels, because it increases work load. In the graph it is seen that this is indeed the case. However, the number of legs correlates greatly with the length of work day. Increasing the number of legs increases duty length, which increases fatigue. Also even though flight duty is long, it may include few short legs and lots of idle time, or couple of very longs legs with few landings and take offs. More legs correlate with amount of work, however, the length of duty may have stronger impact on the fatigue score.

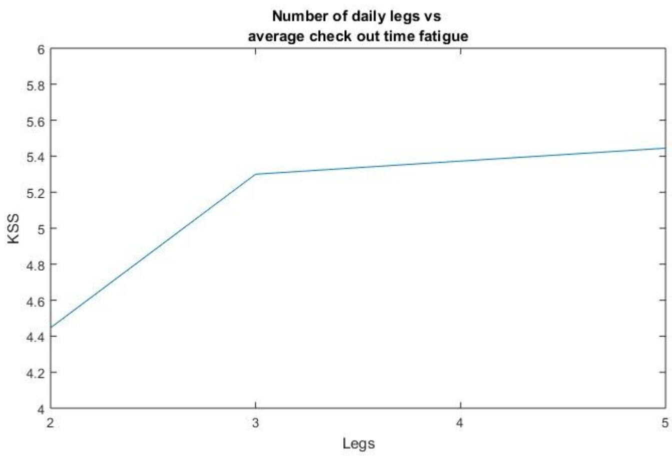

Figure 18. Impact of number of legs to the average after duty fatigue score.

4.3.8. Early consecutive mornings

The effect of consecutive check-in times early in the morning was listed as an important reason for fatigue by the pilots. Due to only 23 data points not equal to zero available for this variable, and most of them equal to 1, it is not possible to conduct meaningful graphical analysis on the effect. From the data available it is possible to calculate that for duties starting at 7:00, or earlier, for mornings with no previous consecutive early starts the mean fatigue score is 3.9 in KSS and with any number of consecutive early mornings 4.8 in KSS. The same figures for check-out are 5.1 and 6.1. So it is clear that consecutive early starts increase fatigue for following days. It is not possible to analyze possible increasing effect due to scarce data, but even one consecutive early start seems to have effect. Based on the data, the disruptive schedule in this case is “late type”, as the effect is clear with duties starting before 7:00.

4.4. Regression Analysis of Model Attributes

4.4.1. Regression model and assumptions

In this section we create a general regression model of the attributes selected for the model: time of day, time spent on duty, cumulative duty, previous rest, cumulative rest and legs flown. This is in order to analyze the attributes and their correlations and the applicability of a regression model. This regression model follows the notation of section 4.2., with addition of $x_{w,t}$ the time worked on current duty at time t, x_r the previous rest time, x_{legs} the number of legs flown on current duty and x_0 the constant of the model. The model is a multiple linear regression model.

$$y(x,t) = x_0 + b_1 x_t + b_2 x_{w,t} + b_3 x_{wcum} + b_4 x_{legs} + b_5 x_r + b_6 x_{rcum}. \qquad (14)$$

Linear regression requires several assumptions (Hoffmann, 2010):

- Linear relationship
- Multivariate normality
- Little multicollinearity

- No auto-correlation
- Homoscedasticity

From the figures in section 4.3. it is evident that the data does not fit very well into the assumption of linear relationship between independent and dependent variables. There is some linearity present, but the data is mostly scattered around, which makes it difficult to fit into regression model. What can be seen is that the relationship is not non-linear, as that would require transformations of the attributes in order to fulfill the linearity assumption. If the relationships are not transformed to linear, information is lost in the regression model.

Testing for normality with Anderson-Darling normality test shows that none of the attributes are from normal distribution. Figure 19 gives histogram plots for the model attributes. The Anderson-Darling test tests for null hypothesis that the data is not from normal distribution and is considered to be an accurate method (Razali et al. 2011). Considering significance tests it is important that the variables are from normal distributions. If this is not the case, the predictions and tests may not be very accurate.

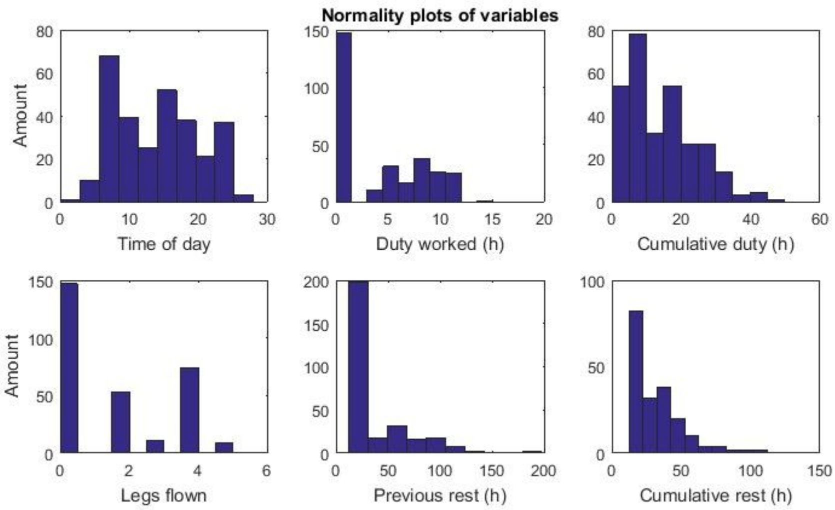

Figure 19. Normality plots for regression model attributes.

Correlation between variables is presented in Table 5. As can be seen, the length of duty, x_{wt}, and number of legs, x_{legs}, have high correlation. This means that they probably contain same information and one of them might be redundant in the model. Same is with x_r and x_{rcum}, but they do contain same information, so it is expected. Other high correlations are between x_t and x_{wt}, and also x_t and x_{legs}. These are due to the fact that the further a day progresses, the more likely it is that the number of legs and the length of duty increase.

Table 5. Correlation of variables.

	x_t	x_{wt}	x_{wcum}	x_{legs}	x_{rcum}	x_r
x_t	1.0000	0.7258	0.4096	0.6896	-0.1516	0.1133
x_{wt}	0.7258	1.0000	0.4045	0.9472	-0.0376	-0.0191
x_{wcum}	0.4096	0.4045	1.0000	0.3811	-0.5726	0.8354
x_{legs}	0.6896	0.9472	0.3811	1.0000	-0.0167	-0.0120
x_{rcum}	-0.1516	-0.0376	-0.5726	-0.0167	1.0000	-0.5659
x_r	0.1133	-0.0191	0.8354	-0.0120	-0.5659	1.0000

The variables are not from normal distributions and examining the scatter plots reveal that linear relationships between independent and dependent variables are not clear. These do not however mean that the regression model would not be applicable.

4.4.2. Regression model results

Table 6 presents the results for the multiple linear regression model for all attributes presented in section 4.4.1.

Table 6. Regression model for all attributes.

Estimated	Coefficients:			
	Estimate	SE	tStat	pValue
x_0	6.0837	0.86686	7.0181	6.2507e-11
x_t	-0.084152	0.034378	-2.4478	0.015466
x_{wt}	0.30606	0.1135	2.6966	0.0077639
x_{wcum}	-0.099839	0.035379	-2.822	0.0053862
x_{legs}	0.16244	0.26044	0.62372	0.53371
x_r	-0.066229	0.050095	-1.3221	0.18805
x_{rcum}	0.033843	0.014811	2.285	0.023644
Root Mean Squared Error: 1.75				
R-Squared: 0.256, Adjusted R-Squared 0.228				

Most of the signs of the coefficients are consistent with expectations. The time of day attribute, x_t, decreases fatigue more as the day goes by, meaning that during morning it has the most effect towards a higher fatigue score. Length of duty increases fatigue as the duty increases. The number of legs has an increasing effect on fatigue and previous rest has a decreasing effect. Both are in line with assumptions that increasing work load increases fatigue while increasing rest time decreases it. Cumulative duty and cumulative rest contradict expectations, as the coefficients are backwards as one would expect. The increase of cumulative work time should have increasing effect on fatigue, and increasing cumulative rest should decrease fatigue, but according to the regression model this is not the case.

The R-squared value describes how well the data fits the model on a scale from 0 to 1, with full correlation at value 1. Because increasing variables increases R-squared value, we use adjusted R-squared to determine the goodness of the regression model. Adjusted R-squared is adjusted to the number of variables in a model; it increases only if the new variable improves the model. In this case the value for adjusted R-squared is low, only 0.228, so the model does not predict very well the fatigue score presented by the data. Root mean square error (RSME) is calculated to be 1.75, which means that on average the estimated value is almost 2 units different on KSS scale.

The p-value for each coefficient tests whether a null hypothesis that the coefficient is zero is true. With significance level of 5%, variables x_{legs} and x_r are not considered to be statistically significant to the model. Especially the variable x_{legs} seems to be insignificant to the model, as it has quite high p-value of 0.53.

4.4.3. Validity of the regression model

The validity of the model is checked with model residuals, which are the error terms in the model and describe random disturbance in the data. Figure 20 shows the residuals in case order plot. It is important that the residuals are uniformly scattered around zero-line, which means that the random error is uniformly distributed across variables and there is no heteroscedasticity. From the figure it can be seen that the residuals seem to be homoscedastic, but there are empty areas on the figure which would indicate heteroscedasticity.

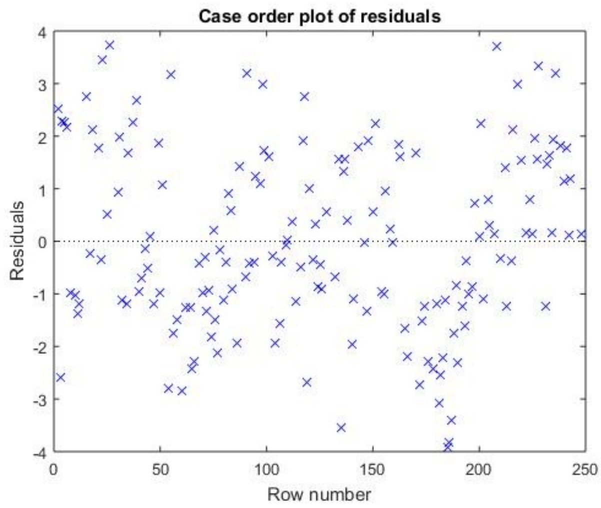

Figure 20. Residuals of the regression model.

Testing for normal distribution with Anderson-Darling test reveals that the residuals are from normal distribution but using Durbinson-Watson test for autocorrelation reveals that the residuals are autocorrelated. Autocorrelation describes time series relationship of data points in a data series, meaning that with autocorrelated data residuals are not independent of each other. A previous value may have effect on some of the following values, as in for example stock markets, where prices are not independent of previous prices.

The low R-squared value implies that the regression model does not predict very well the fatigue score. Analysis of the residuals reveals that a linear regression model may not be suitable for this case, as there is autocorrelation in the model. Another method than regression model should be created to better analyze the fatigue scores with the data collected from pilots.

4.5. Selecting Attributes

4.5.1. Variables with direct effect

Because the regression model created in section 4.4. is not very good at predicting the fatigue score, another method should be used. In section 4.2. is presented a model to calculate fatigue on a discrete scale, starting from predetermined value and simulating the schedule forward. In this section the coefficient values for that model are estimated based on the data.

Coefficients D, E and G must be determined for the model. Using the data gained from the pilots we can calculate the average values for these. Figure 10 and Figure 11 show the effect of duty length to the fatigue score development. We assume in the fatigue model that fatigue increase is linear. Fitting a linear line to the data in Figure 10, and comparing this to the data in Figure 10, we can calculate the hourly drop to be approximately 0.2 units. In 10 hour duty this means a drop of 2 units in KSS. This estimate includes the effect of circadian process, so the coefficient D must be calculated so that on an average duty, the total difference will be 0.2 units/h. Assuming that a workday from 8 to 16 is normal in regard to fatigue, we can calculate the coefficient D so that a workday from 8 to 16 results in increase of 0.2 units/h during work, resulting in difference of 1.6 units in KSS. Recovery is adjusted accordingly, so that on the second day at 8, the fatigue score is on the same level as on the previous day at 8. Due to the regressive nature of recovery process, the starting level of fatigue must be set. We define here that the initial level of fatigue is 70 units, which corresponds to the normal level in KSS. After several rest days crew members are supposed to be rested, so the assumption is valid. The effects of time of day are included in the circadian process.

The value of $D0$ is calculated as follows

$$D0 = \left(d_{dev} + C(16) - C(8)\right) \cdot t_{dev}/h\,, \tag{15}$$

where d_{dev} is the assumed drop in fatigue, t_{dev} the discrete time increment and h the chosen time window, in this case duty length.

The value of E comes from equations (3)-(8). During recovery process $x_w = 0$ so the fatigue score at time t can be presented as

$$H_t = \left(1 - \frac{H_{t-1}}{ul}\right) E + H_{t-1}, \tag{16}$$

which can be presented as the series

$$H_t = \sum_{n=0}^{t-2} (E \cdot c^n) + c^{t-1} \cdot H_1, \tag{17}$$

where

$$c = \left(1 - \frac{E}{ul}\right) \tag{18}$$

with the notation that H_t is the fatigue score which must be achieved with the recovery process and H_1 is the fatigue score from which the recovery process starts. The value of E cannot be calculated analytically, so it must be solved numerically.

To include the effect of consecutive early mornings, a penalty is added, so that

$$H_t = H_T - H_m \cdot x_m, \tag{19}$$

where $x_m \in [0,1]$ is a binary variable based on the presence of consecutive morning shifts starting before or at 7:00, H_T is the fatigue score which is supposed to be reached in normal 8 to 16 roster and $H_m \geq 0$ is the penalty from consecutive morning shifts. The variable x_m is one if current duty and previous duty have early check-in times and zero otherwise.

Table 7 presents the values of D and E when varying the value of circadian coefficient A and alertness scores with the corresponding figures are shown in Figure 21 for standard work day from 8 to 16, with the initial fatigue score of 70.

Table 7. Values for the coefficients D and E by varying A.

Circadian process	**Homeostatic process**	
A	D	E
0	1.00	1.64
5	1.52	2.52
10	2.05	3.55
15	2.57	5.03

The increase of A results in relatively larger increase of coefficient E when compared to coefficient D. This is due to the asymptotic nature of recovery function S', because it limits the recovery at higher levels of alertness and increases it at lower levels.

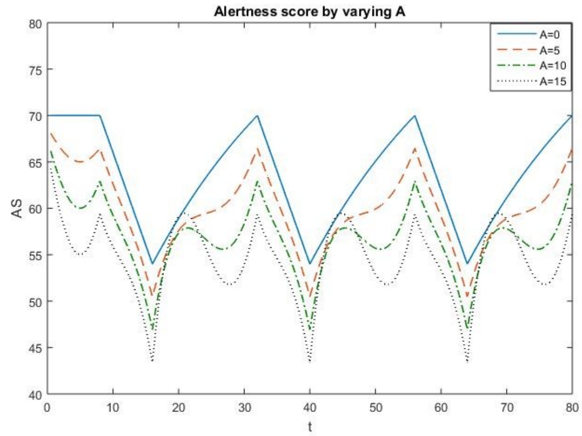

Figure 21. Differences in alertness score by varying A, D and E as in table 7.

As can be seen from Figure 21, the increase of A increases the effect of circadian component. With high values of A this effect is strong when compared to the power of the homeostatic process, as can be seen in the figure when the graphs show more curvature. With the value of 5, the graph in the figure looks realistic, because the AS increases through the recovery period, but the effect of circadian component is still visible. However, the value of A cannot be too low, as the circadian component is important to the fatigue feeling experienced by pilots, according to the data.

Keeping the starting level of the alertness score same, the changing of coefficient A moves the alertness score to a lower level for the whole simulation. If we do not calculate the coefficients D and E for each A, the increase of A increases the effect of circadian process too much when compared to the homeostatic process. Too high A results in AS changing too rapidly depending on the time of day, which diminishes the effect of work and rest periods and puts too much emphasis on the time of day. There must be a balance between circadian and homeostatic processes. Figure 22 presents the alertness scores with varying A but static D and E. The results become too random with increasing values of A for predicting fatigue.

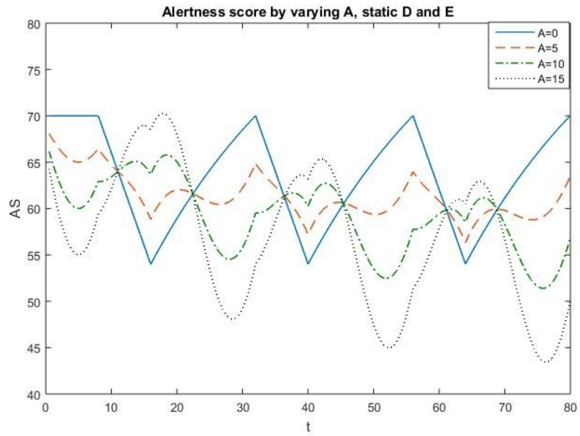

Figure 22. Alertness score with static D and E coefficients but variable A.

4.5.2. Variables with cumulative effect

The cumulative effects of duty and rest times are included in the homeostatic process in equation (7). Increasing the ratio of cumulative duty per cumulative rest increases the coefficient D which results in increased rate of fatigue build up. This is a simple way to measure the work and rest ratio, which should correlate well to the perceived fatigue. If the rest times are very low compared to the duty times, a person will not recover well for the next duty. This is ensured by increasing the rate of fatigue build up, which results in lower alertness scores than without the cumulative factor. According to data in Figure 17, doubling the ratio from 0.3 to 0.6 increases the difference by 1 unit in average. For an average flight duty of 8 hours we can calculate the increase to be $G_r = 0.42$ units in an hour.. Assuming that normal ratio is 8 hours of work for 16 hours of rest, this gives $R_{cum} = 0.5$. Deviations from this will result in increased or decreased cumulative effect to fatigue. We can now define the coefficient G as

$$G = \left(\frac{x_{wcum}}{x_{rcum}} - 0.5\right) \cdot G_r \cdot t_{dev}, \qquad (20)$$

where $G_r \geq 0$, because it can be estimated further.

This considers the effect of overall cumulative workload to fatigue, but the effect of cumulative duty hours as independent variable is not included. The legislation of duty time periods and rest times uses the cumulative duty hours in a given timeframe, such as week or month, as the basis of limiting duty hours. Because the data does not point to the importance of cumulative duty hours in itself, it is left out of the developed model.

The model now developed takes into account the duty lengths, the length of rest periods, the time of day, the ratio of cumulative duty versus cumulative rest and the presence of consecutive early morning shifts. The cumulative effect is set to zero after every rest day between duties. This is done in order to ensure that the R_{cum} measures the ratio between work and rest during consecutive work days. A rest day throws the ratio well over to the rested side, which is not taken into account when calculating the recovery period. The alertness score for the next duty might be too good, so the cumulative factor is reset.

Constants that are estimated from data include the circadian amplitude A, the consecutive early mornings penalty H_m and the cumulative ratio coefficient G_r. As these are only estimates, they can be optimized to increase the accuracy of the model. The estimated fatigue drop of 0.2 units/hour in KSS is considered to be accurate for this model based on the data. The starting level of alertness score is considered to be 70. This initial level is difficult to estimate, if no information is available of the fatigue state of the crew members. The circadian process considers the diurnal variation in check-in times, so further adjustment to the initial value should not be required.

5 Analysis

5.1. General

5.1.1. Model goodness

In order to optimize the parameters in the developed model, only a part of the data is used. Fatigue scores that are only from individual days, which are not part of a series of consecutive work days, are discarded. Series with multiple fatigue scores from work schedule with no rest days between are used. This is done in order to increase the accuracy of the model, because the fatigue scores from separate days may not be accurate, and the model requires initial value for the fatigue level every time simulation is started. Every single fatigue result is therefore subject to error from initial value, so series of fatigue scores are basis for better estimations. After a two day rest a new calculation must be done, because the model does not include accurately the effect of multiple rest days. With these criteria, 22 series of work shifts are chosen. Shortest ones are two days long and longest ones six days long.

Table 8 presents the results for parameters and MSE for the model with all the 22 shifts in the estimation. The optimization is done via brute force method, iterating to one decimal accuracy. Figure 23 presents the MSE for each selected work shift with parameter values from Table 8 and MSE with parameters optimized independently for each work shift, presented in Table 9.

Table 8. Estimated parameter values for the prediction model.

Parameter	A	H_m	G_r	MSE
Value	3.6	12.4	0.0	229.4

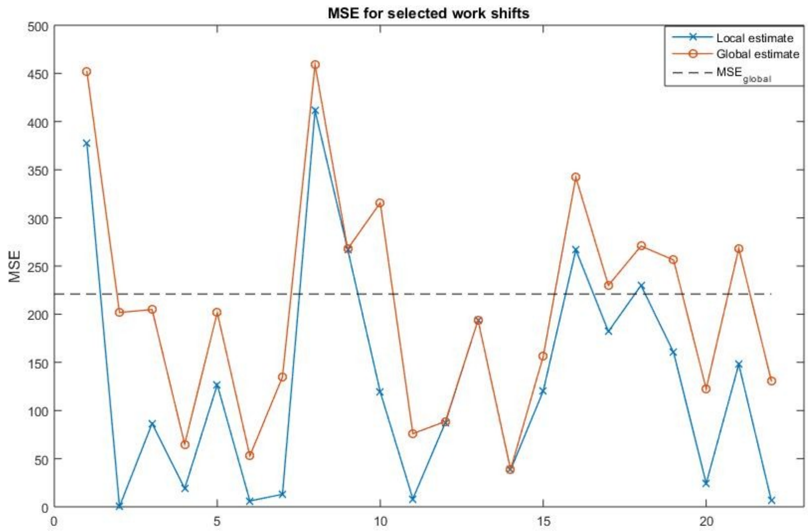

Figure 23. MSE with global and local estimates for parameter optimums.

For some work schedules the difference with local estimate versus the global estimate is very small, whereas some perform much better with local estimates. This is understandable, as the global estimate needs to consider the differences in answers due to personal differences regarding fatigue feelings. Notable is that for several shifts it is possible to achieve very good results, such as for shifts number 6 and 7, and for some, such as shifts 1 and 16, the results are considerably worse even with local estimates. For global estimates the route mean square error, which describes the standard deviation, is $RSME_{global} \approx 15.3$. This means that on average the prediction model estimates almost 1.5 units wrong when transferred to KSS scale.

Table 9 presents the local parameter values and corresponding MSE figures and Figure 24, Figure 25 and Figure 26 show graphical presentations of the parameter estimations with 3D plots. In the 3D plots it is visible that the estimated global parameters are clearly the optimal values in the defined ranges. The MSE value drops smoothly to the minimum value in each plot and there are no local optimums in the figures, only one optimum in each.

Table 9. Parameter estimations and mean square errors for chosen work shifts.

	Parameter			Mean Square Error	
Work shift	A	H_m	G_r	MSE_{local}	MSE_{global}
1	16.0	16.0	0.6	377.5	452.0
2	0.0	0.0	16.3	0.2	201.7
3	13.0	0.0	0.0	86.5	204.8
4	9.0	9.0	1.5	19.0	65.0
5	12.0	0.0	0.0	126.7	201.6
6	3.0	0.0	0.4	6.1	52.9
7	0.0	0.0	0.0	13.0	134.9
8	6.0	30.0	0.0	411.3	459.0
9	3.0	0.0	0.0	267.4	267.8
10	0.0	0.0	0.0	119.3	315.6
11	9.0	14.0	0.8	8.2	75.9
12	3.0	0.0	0.0	87.6	88.7
13	4.0	0.0	0.0	193.8	193.7
14	4.0	0.0	0.0	38.4	38.9
15	4.0	0.0	6.0	120.2	156.1
16	18.0	0.0	0.0	266.8	342.1
17	8.0	0.0	0.0	182.0	229.9
18	8.0	0.0	0.0	229.9	270.8
19	0.0	0.0	0.0	160.3	256.3
20	12.0	0.0	6.8	24.3	122.5
21	14.0	0.0	0.0	147.9	268.2
22	0.0	0.0	0.0	6.3	130.6

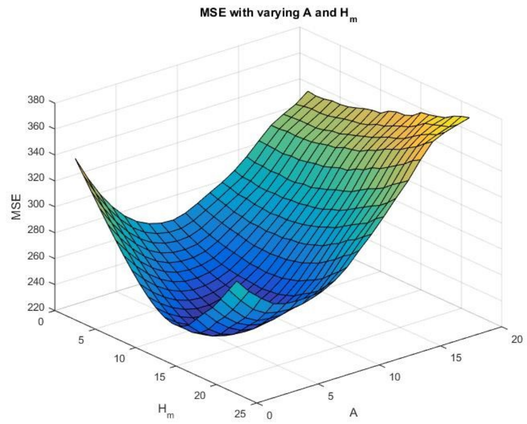

Figure 24. MSE with varying circadian coefficient A and consecutive morning starts coefficient H_m.

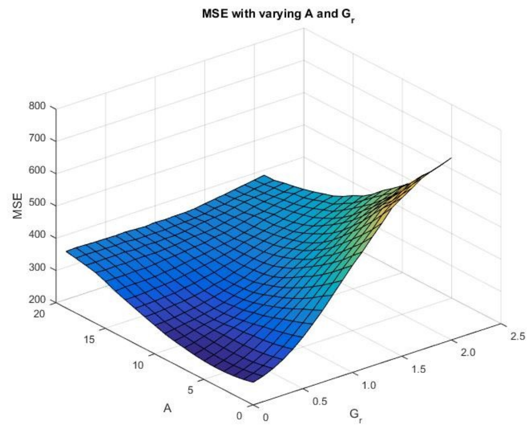

Figure 25. MSE with varying circadian coefficient A and cumulative ratio coefficient G_r.

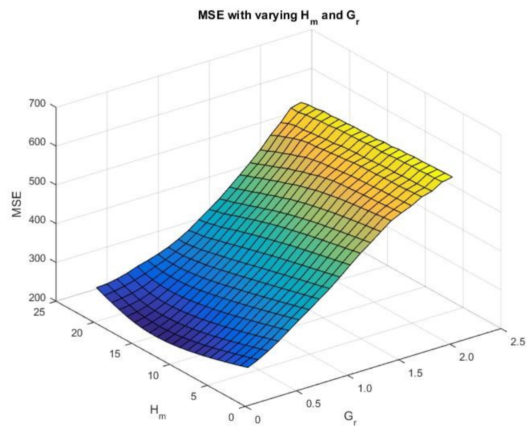

Figure 26. MSE with varying cumulative ratio coefficient G_r and consecutive morning starts coefficient H_m.

Plotting the residuals in Figure 27 and testing with Anderson-Darling test it is found out that they are from normal distribution. With normally distributed residuals the error terms are uniformly distributed, meaning that there is no formulating error in the model, where some information would be lost. Some heteroscedasticity can be seen in the figure, in the form of sinusoidal wave at the center of the picture. The usage of global estimates for all pilots results in some shifts having the error terms on the same side of zero level, because the model calculates uniformly too low or too high fatigue scores for that pilot's preferences. In this regard identifying heteroskedasticity from the figures is questionable.

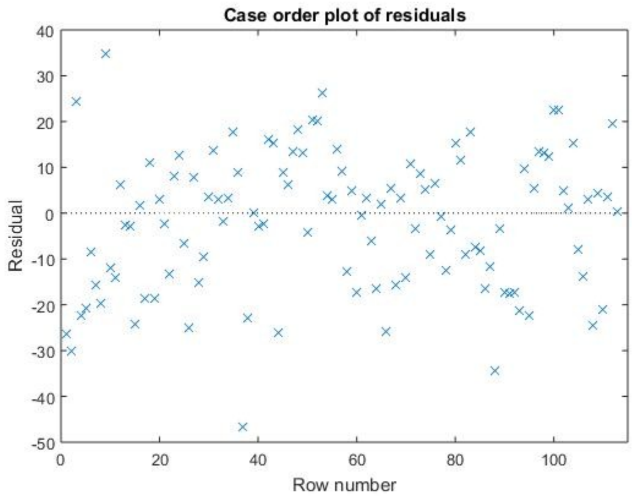

Figure 27. Residuals for the prediction model.

The mean square errors, homoskedasticity and normality of residuals point that the model is applicable to the problem at hand. Though the accuracy is lacking with global estimates, the formulation is correct when regarding available data.

5.1.2. Parameters

The estimated parameters give insight to the data. On the global estimate the circadian coefficient A is not too low or high, meaning that the circadian process is in balance with the homeostatic process. With the local estimates the value varies considerably, describing the differences in data. For some shifts the circadian effect is more important than the homeostatic, hence the high values. In those shifts the length of duty has less effect and the time of day is more important to the fatigue levels. Several reports had same fatigue levels for before and after duty, which is possible to model if the circadian process has considerably large effect when compared to the homeostatic process. Otherwise fatigue increases as duty time increases and the estimate will fit the data poorly.

The coefficient H_m related to early morning starts has high values with both global and local estimates for rosters with relevant duties. This means that for schedules with early morning starts the effect is quite strong. In global estimate the value of $H_m = 12.4$ corresponds to approximately 1.2 units of difference in KSS. With multiple early mornings this seems a reasonable drop in fatigue levels. For shifts with no early mornings this variable is not applicable, so it is mostly zero on local estimates.

The estimate for cumulative ratio G_r is zero for global case and for most local cases. In FTL rules the cumulative effect of work and rest times is considered significant and one can rationalize the importance of it. Estimating from the data the effect is however not clear, because most cases have the variable estimated to be zero. This may be due to the data itself not containing information pointing to the importance of the cumulative effects, or that the proposed meter is not formulated correctly. As the collected fatigue data is mostly from couple of consecutive work days and not from long consecutive duties, the cumulative effect may not be identifiable or present in the available information. In some cases, the impact of cumulative factors is identifiable, because the estimate is non-zero, but that is not enough for global case.

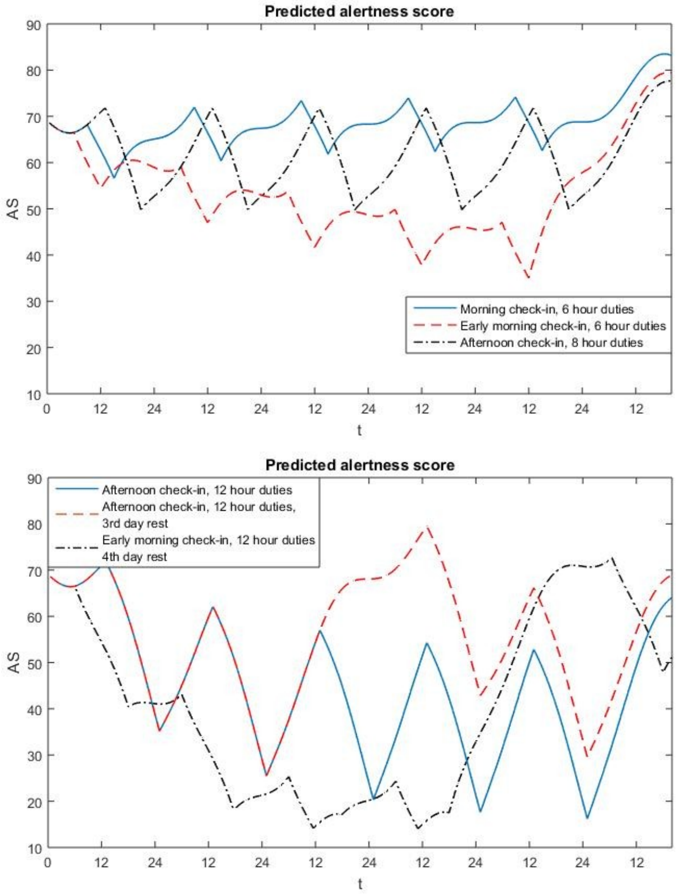

Figure 28. Prediction results for example rosters.

Figure 28 presents results for made-up rosters. The first graph presents short to normal length duties and the second graph maximum length duties for one week roster. The model estimates that the alertness score in normal duties does not drop far below 50 units, and the difference in alertness level at check-in time does not change though duty lengths change. With consecutive early morning check-ins the drop increases and the cumulative effect is clear as week goes by. In the second example of long duties and minimum rest, the alertness level drops after a couple of days close to 20 units at the end of duty, with normal check-ins. With consecutive early morning check-ins the drop is dramatic, and the 3rd day is started with alertness score of 25, which is very low. A rest day reduces the fatigue so that the alertness score returns to normal levels.

Results appear rational and the model seems to capture the problems of early morning starts and long duties, but may over estimate the fatigue effect. Duties appear to be predicted well, because the alertness score does not increase too much with short daily duties, but stays on normal levels. A series of long duties reduces alertness score to low levels and minimum amount of rest is not enough for recovery back to normal levels. Overall the results appear logical, with the effects of duty lengths and time of day showing clearly on the alertness score levels.

5.2. Case examples

In this section the results for some of the shifts used to estimate the parameters are presented. In Figure 29 the first graph is the best case scenario with global estimates, which fits quite well the data points. The second graph is one of the best cases in local estimates, with global and local estimates plotted to the same picture.

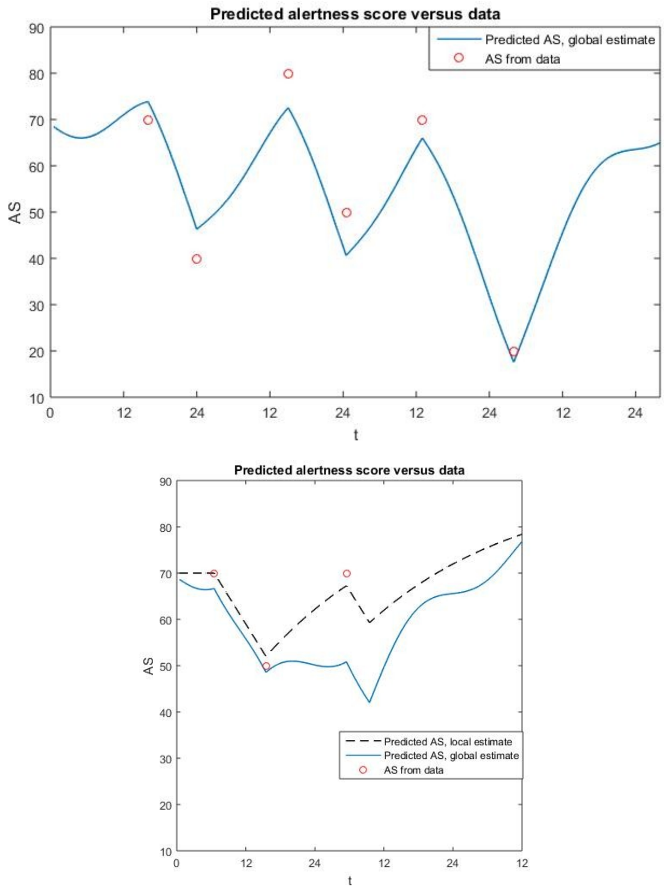

Figure 29. Predicted alertness scores versus data.

As can be seen, the model is able to predict the fatigue especially when local estimates for parameters are used. With global estimates the difference to the local estimates may be high, because the local estimates can identify individual preferences in the data. In the second graph above, the pilot does not feel that the consecutive early morning starts have effect on his fatigue score. The global estimate has the H_m variable estimated as non-zero, so the predicted alertness is much worse than with local estimates.

The differences in pilot answers make it difficult to predict fatigue scores with same parameters for every pilot. Figure 30 presents couple of difficult fatigue scores to model. On the first graph the problem with the data is that between the end of first duty and start of second, there is almost 24 hour rest period, and the next duty starts at 16:30. The created model assumes that this time window is used completely for recovering and the circadian process gives also high alertness for that time. This is quite close to what the pilot has experienced. The second duty lasts only 5 hours, so the model calculates normal levels for alertness score at the time of check-out, which is inconsistent with the data from pilot, as he gives estimate of very low levels of fatigue. The model cannot identify the low level of alertness the pilot is feeling, as the duty is short and circadian timing is favorable as well, but the pilot has reported high levels of fatigue.

In the second graph, the pilot has reported fatigue levels that are mostly same during duties throughout the work week. The prediction model has difficulties calculating the fatigue, because it is assumed that alertness drops during duties as pilots keep working. Problems are also faced with the first alertness scores in both cases, as those are already low even though the pilots have had free days before these shifts. The initial fatigue value should describe the fatigue in normal situation, when the pilot has had adequate rest before duty. As there is no data available concerning the possible initial value, assumptions must be made. In here the assumption is that at the start of work week pilots have normal levels of fatigue. Sleep data confirms that throughout the shifts the pilot in the second case has reported low amounts of sleep on every other day other than the 3rd day, which according to the data has the highest alertness level. The prediction model has no information on the sleep amounts as they cannot be known in advance, so they are not modeled. Sleep however has the highest impact on fatigue levels.

Without the information of actual sleep amounts, the fatigue levels in the second graph are impossible to predict.

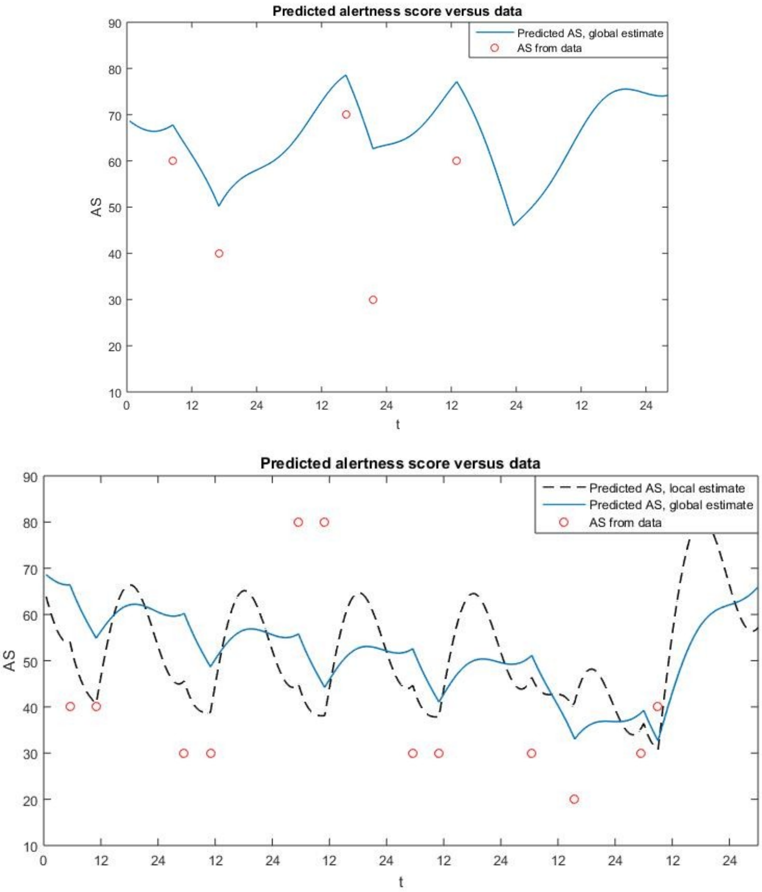

Figure 30. Predicted alertness scores versus data.

The created model is able to find good local estimates for parameters, but not on all cases. The global estimates give satisfying results, but subjective opinions about fatigue make it inaccurate. The results are rational, but when the model is applied to pilot experiences the accuracy is lacking. It is, however, difficult for any model to predict subjective fatigue scores accurately, because the differences in how people perceive fatigue may be high. When comparing to the simple multiple linear regression model, the discrete simulation model gives better results regarding the sensibility of the fatigue scores, because the regression model results may be outside the scale. The RMSE for both regression model and simulation model are on the same scale, so by measuring goodness of fit both are equally good.

6 Conclusions

6.1. Summary

In this thesis, a discrete bio-mathematical model for predicting pilot fatigue in commercial airline operations is developed. The model uses planned work schedules as input and calculates estimates of fatigue levels in discrete time interval. A person is assumed to be either working or resting as indicated by the input and actual sleep is not modeled. Parameters are estimated based on data collected from pilots working on short-haul operations during one month time window in peak season in flight traffic.

Sources for fatigue in the model are time of day, worked duty hours, available rest and presence of consecutive morning shifts. The data used in this thesis does not indicate cumulative work time or cumulative rest time as having significant effect on fatigue, so they do not affect the predicted score; though sleep theory, other models, and aviation regulations assume that the cumulative aspect is important. The effect of work load is also left out off the model, because it is found to be included in the duty length.

The results indicate that it is possible to predict fatigue based on discrete simulation. The model gives rational and feasible results, and when compared to the available fatigue information from pilots, it is able to predict fatigue satisfactorily. The biggest problem in estimating fatigue based on the work schedules is the absence of information regarding the amount of sleep. The actual sleep achieved during previous night is arguably the most important aspect in the following days fatigue levels, but this information is not available in advance. Another issue is that the lengths of rest periods and starting times of duties vary, which requires that the recovery function should either identify when a person is sleeping, or limit the process to prevent excess recovery and alertness levels. In the developed model, time outside of work is defined as recovery period and the parameters are calibrated accordingly. This has the advantage of removing the need of estimating sleep periods out of the model, but raises problems with variable rest times. For increasing rest times the model may over predict alertness levels.

6.2. Model Characteristics

The developed model has several parameters that must be adjusted. The circadian coefficient A defines the effect of the time of day to fatigue. Increasing A increases the effect. Coefficient D defines the rate of fatigue build up and coefficient E the rate of recovery. The coefficients D and E are dependant on each other, and of the coefficient A, and must be defined together. Coefficient H_m defines the penalty for consecutive early mornings and G_r is the cumulative effect. Increasing any of these coefficients other than E will increase fatigue levels. Depending on the disruptive schedule type used, the limit for early morning starts should be set to either 7:00 or 6:00, if following EASA legislation concerning fatigue risk management. Using the limit of 7:00 may include more duties in the disruptive schedule, which results in lower alertness levels due to possible increased number of defined early morning starts.

Utilizing fatigue models requires training to both pilots and the users of the model. Both have to understand how the model works and what it requires from them.

For pilots, the main issue is the adequate amount of sleep. Flight crew should understand that in order for the model predictions to fit, they have to ensure that they acquire enough rest. If they do no utilize rest times well, the real alertness scores will be lower than those predicted by the model. It is important to teach them sleep theory in order to ensure that they understand how sleep and fatigue affect them and how to mitigate the effects. Understanding how to get enough sleep while doing shift work in aviation environment is important for model accuracy.

Users of the model have to understand the limitations and what the results mean. Because the model estimates are only predictions, they must be used with care. The model may be calibrated to fit different types of people or operations, depending on the need. Understanding the parameters, their effect and the interpretation of the results requires training of the users in order to avoid misinterpretations and errors.

6.3. Model Usability

Because the developed model fits only adequately the data, the usability of the results must be questioned. The model may well be used to aid in risk management, but only to certain length. It can be used to identify potential difficult rosters that are then placed under further analysis. In this regard the model fares well, because it aids in decision making, but it cannot be the tool, which the decisions are based on. Further analysis must be done to better rationalize rostering decisions when considering fatigue. It is difficult to justify the usage of the model estimates as a sole basis of decisions, because the results do not reflect the individual fatigue levels.

The differences in how people feel fatigue make it difficult to create an applicable average model that is accurate for everyone. By locally estimating parameters it is possible to achieve much better results, but this requires plenty of data and time. Estimating individual parameters for every pilot is also not practical for an optimization tool aimed for wide spread use in commercial airline operations. There model, however, can be utilized as an average model, because the current

FTL rules are also general rules that do not consider well the individual fatigue perceptions.

For future research, the formulation of the recovery function should be studied further, because the pilot work schedules include changing rest periods, which are difficult to model. Adding a circadian component to the recovery process might improve the accuracy of the recovery modeling. Without prior knowledge of sleep times and amounts, the recovery function has to either rely on probabilistic sleep schedule, assume that time outside of work is used completely for recovery, or find another way to model the recovery process.

References

Achermann, P. & Borbely, A. 2003. Mathematical models of sleep regulation. *Frontiers in Bioscience*. Vol 8, pp. 683-693.

Åkerstedt, T. & Folkard, S. 1997. The Three-process model of alertness and its extension to performance sleep latency and sleep length. *Chronobiology International.* Vol 14, No 2, pp. 115-123.

Åkerstedt, T., Folkard, S. & Portin, C. 2004. Predictions from the three-process model of alertness. *Aviation, Space, and Environmental Medicine*. Vol 75, Supplement 3, pp. A75-A83.

Barnhart, C., Belobabam P. & Odoni, A. 2009. The global airline industry. John Wiley & Sons, Ltd. ISBN 978-0-470-74077-4.

Borbely, A. 1982. A two process model of sleep regulation. *Human Neurobiology*. Vol 1, pp. 195-204.

Branford, K., Lowe, A., Hayward, B. & Cabon, P. 2014. Biomathematical fatigue models. *Guidance Document*. Civil Aviation Safety Authority Australia. Available at: http://www.casa.gov.au/scripts/nc.dll?WCMS:STANDARD::pc=PC_90315 [9.1.2015]

Caldwell, J., Mallis, M., Caldwell, L., Paul, M., Miller, J. & Neri, D. 2009. Fatigue countermeasures in aviation. *Aviation, Space, and Environmental Medicine.* Vol. 80, Issue 1, pp. 29–59.

Civil Aviation Authority UK (CAAUK). 2007. Aircrew fatigue: A review of research undertaken on behalf of the UK Civil Aviation Authority. *CASA Guidance Document.* CAA Paper 2005/04. Available at: http://www.caa.co.uk/application.aspx?catid=33&pagetype=65&appid=11&mode=detail&id=1942 [21.3.2015]

Dawson, D., & Fletcher, A. 2001. A quantitative model of work-related fatigue: Background and definition. *Ergonomics*. Vol 44, No 2, pp. 144-163.

Dawson, D., & McCulloch, K. 2005. Managing fatigue: It's about sleep. *Sleep Medicine Reviews.* Vol 9, pp. 365-380.

Dawson, D., Noy, I., Härmä, M., Åkerstedt, T. & Belenky, G. 2011. Modeling fatigue and the use of fatigue models in work settings. *Accident Analysis and Prevention.* Vol 43, pp. 549-564.

Dean, D., Fletcher, A., Hursh, S., & Klerman, E. 2007. Developing mathematical models of neurobehavioral performance for the "Real World". *Journal of Biological Rhythms*, Vol 22, No. 3, pp. 246-58.

Dinges, D., Graeber, C., Rosekind, M., Samel, A. & Wegmann, H. 1996. Principles and guidelines for duty and rest scheduling in commercial aviation. *NASA Technical Memorandum*, 110404.

EASA. 2012. Commission regulation (EU) No 965/2012. *Official Journal of the European Union.* 5.10.2012.

EASA. 2013. EASA welcomes new flight time limitations rules. Press release 09.10.2013. Available at http://easa.europa.eu/newsroom-and-events/press-releases/easa-welcomes-new-flight-time-limitations-rules. [09.01.2015]

EASA. 2014. Commission regulation (EU) No 83/2014. *Official Journal of the European Union.* 31.1.2014.

ECA 2012. Pilot fatigue barometer. Available at: https://www.eurocockpit.be/stories/20121105/barometer-on-pilot-fatigue [16.9.2015]

Fletcher, A. & Dawson, D. 2001. Field-based validations of a work-related fatigue model based on hours of work. *Transportation Research Part F.* Vol 4, Issue 1, pp. 75-88.

Goode, J. 2003. Are pilots at risk of accidents due to fatigue? *Journal of Safety Research.* Vol 34, Issue 3, pp. 309-313.

Hambour, L. & Majumdar, A. 2011. Validation of fatigue models through operational research. *FRMS forum in Montreal.* Available at: http://www.icao.int/safety/fatiguemanagement/FRMS2011/PRESENTATIONS/Forms/AllItems.aspx [8.1.2015]

Health and Safety Executive (HSE). 2006. The development of a fatigue / risk index for shiftworkers. *Research report 446.* Available at:

http://www.hse.gov.uk/research/rrhtm/rr446.htm [21.3.2015]

Hoffmann, J. 2010. Linear regression analysis: Applications and assumptions. 2nd Edition. Brigham Young University, Provo, USA, 2010.

Hursh, S., Redmond, D., Johnson, M., Thorne, D., Belenky, G., Balkin, T., Storm, W., Miller, J. & Eddy, D. 2004. Fatigue models for applied research in warfighting. *Aviation, Space, and Environmental Medicine*. Vol 75, Supplement 3, pp. A44-A53.

ICAO. 2010. Annex 6 to the convention on international civil aviation, operation of aircraft, part I. Vol 1, 9Th edition, July 2010.

ICAO. 2011. Fatigue risk management systems: Implementation guide for operators, 1st Edition, July 2011. Available at: https://www.iata.org/publications/Pages/frms.aspx [15.3.2015]

ICAO. 2013. Annex 19 to the convention on international civil aviation, operation of aircraft. Vol 1, 1St edition, November 2013.

ICAO. 2014. About ICAO. Available at: http://www.icao.int/about-icao/Pages/default.aspx. [08.01.2015].

InterDynamics. 2014. What you need to know about FAID. Available at: http://www.interdynamics.com/wp-content/uploads/2014/01/WhatYouNeedToKnowAboutFAID.pdf [21.3.2015]

Jeppesen. 2009. Fatigue risk management. *Guidance Document.* Available at: http://ww1.jeppesen.com/documents/aviation/pdfs/Fatigue_2009-10_Final_II.pdf [21.3.2015]

Jewett, M. & Kronauer, R. 1999. Interactive mathematical models of subjective alertness and cognitive throughput in humans. *Journal of Biological Rhythms.* Vol 14, No 6, pp. 588-597.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T. & Fukasawa, K. 2006. Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology*. Vol 117, pp. 1574-1581.

Kandelaars, K., Dorrian, J., Fletcher, A., Roach, G. & Dawson, D. 2005. A review of bio-mathematical fatigue models: Where to from here? *Sixth International Conference on Fatigue Management in Transportation.* Seattle, Washington, USA. 11-15 Sep 2015.

Moore-Ede, M., Heitmann, A., Guttkuhn, R., Trutschel, U., Aguirre, A., & Croke, D. 2004. Circadian alertness simulator for fatigue risk assessment in transportation: Application to reduce frequency and severity of truck accidents. *Aviation, Space, and Environmental Medicine.* Vol 75, No. 3, Section II, pp. A107-A118.

Powell, D., Spencer, M., Holland, D., Broadbent, E. & Petrie, K. 2007. Pilot fatigue in short haul operations: Effects of number of sectors, duty length, and time of day.

Aviation, Space, and Environmental Medicine. Vol. 78, pp. 698-701.

Rangan, S., & Van Dongen, H.P.A. 2013. Quantifying fatigue risk in model-based fatigue risk management. *Aviation, Space, and Environmental Medicine*. Vol. 84, pp. 155-157.

Razali, N. & Wah, Y. 2011. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*. Vol 2, No 1, pp. 21-33.

Roach, G., Fletcher, A. & Dawson, D. 2004. A model to predict work-related fatigue based on hours of work. *Aviation, Space, and Environmental Medicine.* Vol 75, Supplement 1, pp. A61-A69.

Roma, P., Hursh, S., Mead, A. & Nesthus, T. 2012. Flight attendant work/rest patterns, alertness, and performance assessment: Field validation of biomathematical fatigue modeling. *Report No. DOT/FAA/AM-12/12.* Washington, DC: Office of Aerospace Medicine.

Samn, S.W., & Perelli, L.P. (1982). Estimating aircrew fatigue: A technique with implications to airlift operations. *Technical Report No. SAM-TR-82-21.*

Steiner, S., Fakles, D. & Gradisar, T. 2012. Problems of crew fatigue management in airline operations. *Internal Conference on Traffic and Transport Engineering*. Belgrade, November 29-30, 2012.

Trafi [a]. 2015. Aviation Organization. Available at: http://www.trafi.fi/en/aviation [16.9.2015]

Trafi [b]. 2015. Implementation of EASA FTL. Available at: http://www.trafi.fi/ilmailu/saadokset/easa/easa_ftl_kayttoonotto [8.12.2015]

Van Dongen, H. & Belenky, G. 2011. Model based fatigue risk management. *Handbook of Operator Fatigue*. Aldershot: Ashgate Publishing.